

PHASE TRANSITIONS IN A SPATIAL COALESCENT

BY NIC FREEMAN*

Oxford University

We construct a natural extension of the Λ -coalescent to a spatial continuum, and analyse its behaviour. Like the Λ -coalescent, the individuals in our model can be separated into (i) a dust component and (ii) large blocks of coalesced individuals. We identify a five phase system, where our phases are defined according to changes in the qualitative behaviour of the dust and large blocks. We completely classify the phase behaviour, and obtain necessary and sufficient conditions for the model to come down from infinity.

We believe that two of our phases are new to Λ -coalescent theory, and reflect the incorporation of space into our model. Firstly, our semicritical phase sees a null but non-empty set of dust. In this phase the dust becomes a random fractal, of a type which is closely related to iterated function systems. Secondly, our model has a critical phase, in which the total number of blocks becomes a.s. finite at a deterministic, positive time.

1. Introduction.

1.1. *Λ -coalescents.* The theory of coalescent processes began with [Kingman \[1982\]](#), as a means of analysing the relationships between genes carried within a non-spatial population. Since then, coalescents have become an important part of mathematical biology, usually as duals to population models. Loosely speaking, coalescent processes are systems of particles, which start out separated, and merge together into ever growing clusters over time. A recent survey coalescent theory, and its other applications, can be found in [Berestycki \[2009\]](#).

Kingman's coalescent was generalised, independently but in the same spirit, by [Donnelly and Kurtz \[1999\]](#), [Pitman \[1999\]](#) and [Sagitov \[1999\]](#). The resulting process became known as the Λ -coalescent, which we now describe. Let \mathcal{P} denote the set of partitions of \mathbb{N} , and let \mathcal{P}_n denote the set of partitions of $\{1, 2, \dots, n\}$. Let π_n denote the natural restriction map $\pi_n : \mathcal{P} \rightarrow \mathcal{P}_n$, which is defined by simply removing all elements $m \in \mathbb{N} \setminus \{1, \dots, n\}$ from a partition of \mathbb{N} . For example, $\pi_3(\{(1, 2, 6), (3, 5), (4), \dots\}) = \{(1, 2), (3)\}$. Let $1^{\mathbb{N}}$ denote the partition of \mathbb{N} into singletons.

*Supported by EPSRC Studentship MATH0906.

Key words: Λ -coalescent, Spatial Λ -coalescent, Phase transition, Galton-Watson, varying environment. *MSC 2010 Subject Classification:* 60J25 60J85 92D25.

If A is a partition of some set, then each element of A is known as a *block*. If $A = \{A_1, A_2, \dots, A_b\}$ and $I \subseteq \{1, \dots, b\}$ then the partition obtained from A by *merging* $\{A_i; i \in I\}$ is given by $\{A_i; i \notin I\} \cup \{\cup_{i \in I} A_i\}$.

DEFINITION 1.1. *Let Λ be a finite measure on $(0, 1]$. The Λ -coalescent is the unique \mathcal{P} -valued Markov process $(\Pi_t)_{t \geq 0}$ such that, for all $n \in \mathbb{N}$, $\Pi_t^{(n)} = \pi_n(\Pi_t)$ is a \mathcal{P}_n -valued Markov chain with initial state $\pi_n(1^{\mathbb{N}})$ and the following dynamics: Whenever $\Pi_t^{(n)}$ is a partition consisting of b blocks, the rate at which any k -tuple of blocks merges is*

$$(1.1) \quad \lambda_{b,k} = \int_0^1 x^{k-2} (1-x)^{b-k} \Lambda(dx),$$

independently of all other k -tuples.

The formulation of Definition 1.1 is due to Pitman [1999], and the reader is referred there for further detail. If $\Lambda(\{0\}) = 0$, then the formula (1.1) is more intuitively written as $\lambda_{b,k} = \int_0^\infty x^k (1-x)^{b-k} x^{-2} \Lambda(dx)$. The term $\nu(dx) = x^{-2} \Lambda(dx)$ corresponds to a measure controlling the rate at which a proportion $x \in (0, 1]$ of the blocks currently present merge to form a new block. The remaining ‘binomial’ term $x^k (1-x)^{b-k}$ says that, of the first b blocks, each block chooses independently whether to become part of the new block, or remain alone (with probabilities x and $1-x$ respectively).

Kingman’s coalescent corresponds to the special case where Λ is a point-mass at 0. In this case, as is readily seen from (1.1), each merger involves only a pair of blocks.

For us, it is important to understand the character of the process; the Λ -coalescent is a \mathcal{P} -valued Markov process in which, at random points in time, a selection of the currently present blocks merge into a single block.

In this article we do not consider Ξ -coalescents, which are \mathcal{P} -valued coalescent processes in which more than one new block may be created at the same instant of time. The family of Ξ -coalescents, which further generalize Kingman’s coalescent, were introduced by Schweinsberg [2000b] and Möhle and Sagitov [2001]. They have been studied in, for example, Limic [2010].

It is usual to equip the set of blocks of the Λ -coalescent with a genealogy, defined as follows. We say a block A , created in a merger at time t from blocks $A = \cup_{i \in I} A_i$, is a parent with children $\{A_i; i \in I\}$. Thus each realization of $t \mapsto \Pi_t$ gives rise to a natural tree structure on the set $\cup_t \Pi_t$, and on the restrictions $\cup_t \Pi_t^{(n)}$.

Let $|A|$ denote the cardinality of A when A is a finite set, and set $|A| = \infty$ otherwise. The initial state of Π_t is $\Pi_0 = 1^{\mathbb{N}}$, and clearly $|1^{\mathbb{N}}| = \infty$. Pitman

[1999] proved that, if there is no possibility of a single reproduction event merging the whole population (i.e. $\Lambda(\{1\}) = 0$), then, almost surely, either (1) for all $t > 0$, $|\Pi_t| = \infty$, or (2) for all $t > 0$, $|\Pi_t| < \infty$.

DEFINITION 1.2. *If $|\Pi_t| < \infty$ a.s. for all $t > 0$ then we say that the Λ -coalescent comes down from infinity.*

Schweinsberg [2000a] gave a necessary and sufficient condition, in terms of the $\lambda_{b,k}$, for the Λ -coalescent to come from infinity. In view of Pitman's result, it is natural to think of the Λ -coalescent as experiencing a phase transition, and of Schweinsberg's result as identifying the phase boundary.

When the Λ -coalescent does come down from infinity, considerable effort has been devoted to establishing the asymptotic growth of $|\Pi_t|$ as $t \downarrow 0$. In general such questions do not have an easy answer, but for the special case of Kingman's coalescent it is not hard to see that $|\Pi_t| \asymp t^{-1}$ as $t \downarrow 0$ (for example, as in Section 2.1.2 of Berestycki 2009).

The first example of a (non-Kingman) Λ -coalescent appeared in Bolthausen and Sznitman [1998], and corresponds to the case where Λ is the uniform measure on $[0, 1]$. Bertoin and Le Gall [2000] discovered a correspondence between the genealogy of the Bolthausen-Sznitman coalescent and the genealogy of the continuous state branching process (CSBP) considered in Neveu [1992]. This correspondence was extended to more general CSBPs and β -coalescents by Birkner et al. [2005]. Recall that β -coalescents are the subclass of Λ -coalescents in which Λ has the $\beta(2 - \alpha, \alpha)$ distribution, for $\alpha \in (0, 2)$. The $\beta(1, 1)$ distribution is uniform on $[0, 1]$, so the case $\alpha = 1$ is the Bolthausen-Sznitman coalescent. The connection to CSBPs was used in Berestycki et al. [2008] to establish the asymptotic behaviour of $|\Pi_t|$ for β -coalescents. It turns out that β -coalescents come down from infinity if and only if $\alpha \in (1, 2)$, and in this case $|\Pi_t| \asymp t^{1/(\alpha-1)}$.

Returning to the general case, Bertoin and Le Gall [2006] gave a condition (equivalent to that of Schweinsberg [2000a]) for the Λ -coalescent to come down from infinity, in terms of the connection to CSBPs. Berestycki et al. [2010] exhibit a function $v(t)$ such that, for a general Λ -coalescent, $|\Pi_t| \asymp v(t)$. Their proofs use a martingale method, rather than CSBPs, but the function $v(t)$ bears a close resemblance to the formulation of the coming down from infinity condition in Bertoin and Le Gall [2006].

In summary, branching processes are an important tool for studying coalescents. We continue the theme in this article, since our own model has a close connection to a family of Galton-Watson processes in varying environments. The connection will be described in Section 1.5.

1.2. *Spatial versions of Λ -coalescents.* We have mentioned that our model is a spatial version of the Λ -coalescent. A mean-field version of the Λ -coalescent has already appeared, in [Limic and Sturm \[2006\]](#), building on the mean-field version of Kingman’s coalescent from [Greven et al. \[2005\]](#). The model from [Limic and Sturm \[2006\]](#) is referred to in the literature as ‘the spatial Λ -coalescent’. We will usually refer to our own generalisation as ‘our model’ or as ‘the segregated Λ -coalescent’. The use of the term ‘segregated’ corresponds to our choice of geographical space (defined in [Section 2.1](#)), which we call a *segregated space*.

Let us give a brief description of the model considered in [Limic and Sturm \[2006\]](#). Let G be a finite graph. Initially, at each vertex $g \in G$ we have a copy of $1^{\mathbb{N}}$ (the partition of \mathbb{N} into singletons). As time passes two mechanisms operate simultaneously. Firstly, each block (from any vertex) has an independent spatial motion, causing it to move between the vertices of G . Additionally, at each vertex $g \in G$ we run mutually independent Λ -coalescents, operating on all the blocks which are currently at the given vertex. Thus, blocks wander freely around G , but are intermittently coagulated with other blocks which happen to be at the same vertex at the same time. [Limic and Sturm \[2006\]](#) obtained conditions analogous to those of [Schweinsberg \[2000a\]](#) for the spatial Λ -coalescent to come down from infinity.

However, [Angel et al. \[2010\]](#) show that, in the same model but with G countably infinite and of bounded degree, the resulting process does not come down from infinity (counting blocks from all sites of G together). In contrast, our own model can come down from infinity. [Angel et al. \[2010\]](#) also obtain asymptotic results on the behaviour of their coalescent, in the Kingman and β -coalescent cases, for the behaviour of a large (but finite) number of blocks.

Our model is very different to that of [Angel et al. \[2010\]](#), and although we delay a description of our model until [Section 1.4](#), we can outline a few big differences here. The geographical space in our model is a spatial continuum, not a graph, and we permit only a single block to occupy each spatial location at any one time. The blocks in our model will not move in space, except during mergers. Further, when a merger occurs in our model it will involve all the blocks from within a non-null proportion of the geographical space. Thus, the multiple mergers in the two models are ‘multiple’ for quite different reasons, and it is natural to expect different behaviour.

1.3. *Spatial Λ -Fleming-Viot processes.* In another vein, attempts were made to further generalise population models and incorporate the effect of space (in the geographical sense). For our purposes, the transition towards

spatial models begins with the Λ -Fleming-Viot process, which is a dual to the Λ -coalescent, constructed in Bertoin and Le Gall [2003] (and implicit in Donnelly and Kurtz 1999). The Λ -Fleming-Viot process is a natural generalisation of the Fleming-Viot process, which is itself dual to Kingman's coalescent.

The spatial Λ -Fleming-Viot process (SAFV) was introduced in Etheridge [2008] as a spatial extension of the Λ -Fleming-Viot process. The SAFV is an infinite system (one at each site of \mathbb{R}^d) of interacting Λ -Fleming-Viot processes. We describe a simplified version of the SAFV process in Definition 1.5.

REMARK 1.3. *The terminology 'the SAFV process' is usually reserved for the process considered in Barton et al. [2010a], but both Etheridge [2008] and Barton et al. [2010a] stress that the SAFV is only one illustration of a much more general theme, namely population models in a spatial continuum with reproduction controlled by a Poisson point process. We refer to such a process as 'a SAFV process'. The variety of possibilities is demonstrated by Berestycki et al. [2009], Barton et al. [2010b] and Etheridge and Véber [2011].*

Since the dual of the Λ -Fleming-Viot process is the Λ -coalescent, and the SAFV is a spatial version of the Λ -Fleming-Viot process, the dual to a SAFV process behaves like a spatial version of the Λ -coalescent. We will give a rudimentary example of such behaviour with Definition 1.5. Our own model (which is defined in Section 1.4) provides a much better example, because our results show that the coalescent behaviour can be greatly enriched by the introduction of space.

REMARK 1.4. *As well as their construction of the SAFV process, Barton et al. [2010a] study the space-time rescaling of their dual process, and establish a wide range of behaviour, with connections to both Kingman and more general Λ -coalescents.*

The SAFV process has multiple individuals at each site, but has a natural simplification in which each site carries precisely one individual at any one time. We will now describe this version of the process.

DEFINITION 1.5 (simple version of SAFV). *Let π be a Poisson point process with points $(t, x, r) \in \mathbb{R} \times \mathbb{R}^d \times (0, \infty)$ and intensity measure*

$$dt \otimes dx \otimes \mu(dr).$$

Here dt and dx denote Lebesgue measure. The dynamics are as follows. At any one point in time, each site $y \in \mathbb{R}^d$ carries a single individual. When $(t, x, r) \in \pi$, the individual at x at time $t-$ becomes a parent, all individuals in $B_r(x) = \{y \in \mathbb{R}^d; |x - y| < r\}$ at time $t-$ are killed, and the children of x instantaneously colonise $B_r(x)$ at time t .

We refer to (t, x, r) as a *reproduction event* occurring in $B_r(x)$ at time t , and to x as the parent of the event.

As a consequence of the construction in Barton et al. [2010a], the process of Definition 1.5 exists (with an appropriate choice of state space) under the condition

$$(1.2) \quad \int r^d \mu(dr) < \infty.$$

Let us take a moment to understand (1.2). Recall that each site of \mathbb{R}^d is occupied, at any time, by precisely one individual. The individual at y is affected by reproduction events at rate

$$\int_{\mathbb{R}^d} \int_0^\infty \mathbb{1}\{y \in B_r(x)\} \mu(dr) dx = C_d \int_0^\infty r^d \mu(dr),$$

where C_d is the d -dimensional volume of $B_1(0)$. So (1.2) says precisely that the reproduction mechanism of Definition 1.5 affects each point (equivalently, each individual) at finite rate.

Consider tracing the ancestral lineages of this process. Informally, this means choosing some individual and tracing back the line of parents which gave rise to that individual. Since each point carries only a single individual, it is equivalent to tracing the spatial location of the line of parents.

By (1.2), reproduction events affect individuals at finite rate, so it is also the case that a single ancestral lineage jumps to a new individual at finite rate. The spatial homogeneity of π implies that the parent site, to which the lineage jumps, has the same displacement (in distribution) from the original position of each jump. A little extra work shows that the ancestral lineage of a single individual is a compound Poisson process.

Thus, the ancestral lineages form a system of interacting compound Poisson processes. Since a single reproduction event affects multiple sites, this system of ancestral lineages sees multiple mergers. It follows from properties of the Poisson point process that at most one merger occurs at any (random) instant in time. As claimed, we see a spatial equivalent of the Λ -coalescent.

The reader might wonder why, for the process in Definition 1.5, we have to restrict to compound Poisson process as ancestral lineages, and not use general Lévy processes. We refer the reader to Barton et al. [2010a] for a proper

discussion, but it boils down to the issue of actually *defining* a suitable system of coalescing Lévy processes. The difficulty comes from an apparent incompatibility of the compensation mechanism of the Lévy processes and the driving Poisson point process π .

It would not be fair to claim that we have overcome this difficulty in our model. Rather, we choose our state space and reproduction mechanism in such a way as our ancestral lineages can jump at infinite rate, but without the need for Lévy process style compensation.

The SAFV process in [Barton et al. \[2010a\]](#) uses a non-trivial ‘ L^p type’ state space, which loses track of any behaviour occurring on null sets. Since a spatially null set could contain infinitely many individuals, for our purposes we must keep track of null sets. This is achieved by formulating our model as a stochastic flow on the geographical space.

Coalescent processes and stochastic flows are two different methods of representing the coalescence of particles over time. It is natural to expect connections between the two, in fact a correspondence between (non-spatial) Λ -coalescents and a family of stochastic flows was established in [Bertoin and Le Gall \[2003\]](#).

DEFINITION 1.6. *Let $(X_{s,t})_{s < t}$ be a stochastic flow on some space K . We say X comes down from infinity if $X_{0,t}(K) = \{X_{0,t}(x); x \in K\}$ is a.s. finite for all $t > 0$.*

Studying when, and at what rate, stochastic flows come down from infinity is not new, though many such works would not have considered themselves as part of coalescent theory. For example, it is well known, due to [Arratia \[1979\]](#), that the Arratia flow comes down from infinity, on a compact subset of \mathbb{R}^d .

Due to (1.2), the model of Definition 1.5 does not come down from infinity. To see this, note that (1.2) implies that the probability of the individual at $x \in \mathbb{R}^d$ being affected by no reproduction events during $[0, t]$ is positive. Let $p_t \in (0, 1]$ be this probability. Fix $R > 0$ and let A be the subset of $B_R(0)$ consisting of sites which have not been affected by reproduction events. Let $\text{vol}(A)$ denote the d -dimensional volume of A . By Fubini’s theorem,

$$(1.3) \quad \mathbb{E}[\text{vol}(A)] = \int_{B_R(0)} \mathbb{E}[\mathbb{1}\{x \in A\}] \, dx = \int_{B_R(0)} p_t \, dx = C_d R^d p_t > 0.$$

Thus A is non-null (and hence uncountable) with positive probability. In the language of coalescents, with positive probability the dust contains a positive fraction of the population. A similar argument may be applied to

the SLFV process of [Barton et al. \[2010a\]](#), in its full generality, obtaining an equivalent result.

1.4. Our model. We now move on to defining our own model and stating the results which are proved in later sections. We introduce our model in the special case where the geographical space K is the Cantor set, and delay full generality until Section 2. We use the symbol \uplus to denote a disjoint union of sets.

It is well known that the 2-part Cantor set K is the unique non-empty compact subset of $[0, 1]$ which satisfies $F_1(K) \uplus F_2(K) = K$, where $F_1(x) = x/3$ and $F_2(x) = 2/3 + x/3$. We call the set

$$K_{i_1 \dots i_n} = F_{i_1} \circ \dots \circ F_{i_n}(K)$$

an n -complex of K , or, when n is not specified, a complex of K . If $K_w \subseteq K_{w'}$ then we say $K_{w'}$ is a subcomplex of K_w . If $w = i_1 \dots i_n$, where $i_j \in \{1, 2\}$, then we write $|w| = n$. With mild abuse of notation, define $K = K_\emptyset$, where \emptyset is the empty set and set $|\emptyset| = 0$. As usual, K is the disjoint union of its n -complexes:

$$(1.4) \quad K = \biguplus_{i_1 \dots i_n \in \{1, 2\}^n} K_{i_1 \dots i_n}.$$

In order to illustrate sufficient generality we will need the $|S|$ -part Cantor set, where $|S| \in \{2, 3, 4, \dots\}$. This is the unique non-empty compact subset of $[0, 1]$ which satisfies $K = \biguplus_{i=1}^{|S|} F_i(K)$ where $F_i(x) = \frac{1}{2|S|-1} (2i - 2 + x)$. The 1-complexes $K_i = F_i(K)$ consist of $|S|$ copies of K , each shrunk by a factor of $\frac{1}{2|S|-1}$, evenly spaced across $[0, 1]$. The notions of complex and subcomplex are extended in the obvious manner, and the obvious analogue of (1.4) holds.

We write $|S|$, instead of a single symbol, so as our present notation agrees with the more general setup in Section 2.1 (note that $S = \{1, 2, \dots, |S|\}$).

REMARK 1.7. *Our model is not restricted to totally disconnected scenarios like the Cantor set, and we are able to deal with situations where the complexes touch (that is, where $\overline{K_w} \cap \overline{K_{w'}} \neq \emptyset$ and $K_w \cap K_{w'} = \emptyset$). We define our geographical space K in full generality in Section 2.1.*

Our model is parametrised by $|S|$ and a sequence $(r_n)_{n=0}^\infty \subseteq [0, \infty)$. To avoid triviality we require that $|S| \geq 2$ and $r_n > 0$ for some n . The model is formulated as a stochastic flow $X_{s,t} : K \rightarrow K$ where $-\infty < s < t < \infty$. We delay a fully rigorous definition of the flow $X = (X_{s,t})$ until Section 2.3, and for now content ourselves with the following.

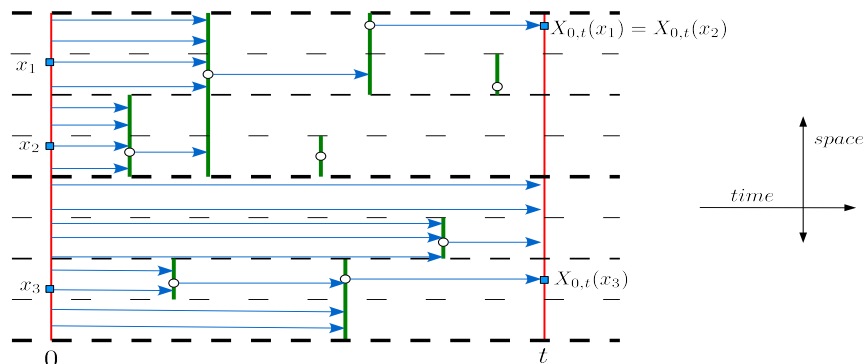


FIG 1. The complexes of the geographical space with $|S| = 2$ are shown down to level 3, with dotted lines. In this realisation there are no reproduction events occurring in complexes of level 4 and above. The movement of some sample particles over $[0, t]$ is indicated by arrows, with the start/end-points from x_1, x_2, x_3 marked with square dots. Reproduction events are shown as thick vertical lines, with parents as circular dots.

DEFINITION 1.8 (informal). To each complex K_w , associate an independent exponential clock of rate $r_{|w|}$. If the clock for K_w rings at time t , a point p is sampled uniformly from K_w . Any particles in the flow which are in K_w at time t — jump to p at time t . In between times at which the clocks ring, the particles in the flow do not move.

We write the resulting flow as $(X_{s,t})_{s < t}$, where $X_{s,t} : K \rightarrow K$.

When the clock for K_w rings, we say a reproduction event has occurred in K_w at time t with parent p . Often we will shorten this to ‘level n reproduction event’, where $|w| = n$. Figure 1 gives a graphical demonstration of Definition 1.8 in the case of $X_{0,t}$ when $|S| = 2$.

Definition 1.8 makes rigorous sense if, for example, $\sum_0^\infty |S|^n r_n < \infty$. This condition characterises the case where the total rate of all the exponential clocks is finite, so as only finitely many reproduction events occur during any bounded interval $[s, t]$.

Under Definition 1.8, a single point of K is affected by reproduction events at rate $\sum_0^\infty r_n$. By comparison with (1.2) and Barton et al. [2010a], we might at first assume that the condition $\sum_n r_n < \infty$ was necessary for the models existence. This is not so, in fact in Section 2.3 we give a direct definition of $(X_{s,t})$ without placing any assumptions on $(r_n) \subseteq [0, \infty)$ or $|S| \in \mathbb{N} \setminus \{1\}$.

The key ingredient in our existence argument is the self similar structure of the geographical space, and the fact that, for two reproduction events,

either one occurs inside the subset of K in which the other occurs, or they occur in disjoint subsets of K . Crucially, large events can overwrite the effect of preceding smaller events.

We now record a full statement of the existence theorem, with the understanding that a mathematically rigorous definition of the flow X will be given in Section 2.3.

Let D_K denote the Euclidean metric on $K \subseteq [0, 1]$. Let \mathbb{M} be the metric space of measurable functions mapping K into K , equipped with the metric $\|f, g\|_\infty = \sup\{D_K(f(x), g(x)); x, y \in K\}$.

THEOREM 1. *For each $s < t$, $X_{s,t}$ is an \mathbb{M} -valued random variable.*

- *For all $s < t < v$, $X_{s,v} = X_{t,v} \circ X_{s,t}$.*
- *For all $s < t < u < v$, $X_{u,v}$ and $X_{s,t}$ are independent.*
- *For all $t_1 - s_1 = t_2 - s_2$, X_{s_1,t_1} and X_{s_2,t_2} are identically distributed.*

The formula $X_{s,u} = X_{t,u} \circ X_{s,t}$, is known as the flow property, and shows that the population which our model describes has a consistent genealogical structure.

We delay a discussion of measurability/continuity of X until after we have defined our model in full generality. See Remarks 2.12-2.14.

The segregation of K into complexes, coupled with the fact that reproduction events occur only in the complexes, means that the probabilistic structure of our model is significantly more tractable than the similar model described in Definition 1.5. Note in particular that, in our model, a large number of small reproduction events cannot cause a particle to ‘creep’ far away from its starting point. A further illustration of tractability is found in a connection to Galton-Watson processes in varying environments, which we now describe.

1.5. The connection to GWVEs. Let $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. A Galton-Watson process in a varying environment (GWVE) is a classical Galton-Watson process, with the modification that the offspring distribution of an individual may depend on its generation number.

To be precise, a GWVE $(B_n)_{n \in \mathbb{N}_0}$ is specified by an initial random value $B_0 \in \mathbb{N}_0$ and a sequence $(\mathcal{G}^{(n)})_{n \in \mathbb{N}}$ of \mathbb{N}_0 -valued random variables, called offspring distributions. Let $(\mathcal{G}_i^{(n)})_{i=1}^\infty$ is a sequence of independent copies of $\mathcal{G}^{(n)}$. The GWVE is defined by some initial (independent) random variable $B_0 \in \mathbb{N}_0$, and the inductive relation $B_{n+1} = \sum_{i=1}^{B_n} \mathcal{G}_i^{(n)}$. The classical Galton-Watson process corresponds to the case where $\mathcal{G}^{(n)}$ has the same distribution, for all n .

The connection to our model is as follows. For each $t > 0$ and $n \geq 0$ define

$$\mathcal{B}_n^t = \{K_w; |w| = n, \text{ and for no complexes } K_{w'} \supseteq K_w \text{ did the clock for } K_{w'} \text{ ring during } (0, t]\}.$$

Set $B_n^t = |\mathcal{B}_n^t|$. Then, for each fixed $t > 0$, $(B_n^t)_{n \geq 1}$ is a GWVE. To see this, note first that $e^{-r_n t}$ is the probability that K_w , where $|w| = n$, does not see its clock ring during $(0, t]$. If $w \in \mathcal{B}_n^t$ and $|w| = n$, then the (conditional) probability that $wi \in \mathcal{B}_{n+1}^t$ is just $e^{-r_{n+1} t}$. The clocks corresponding to K_{wi} and K_{wj} are independent if $i \neq j$, thus the offspring distribution of $w \in \mathcal{B}_n^t$ is Binomial with $|S|$ trials and success probability $e^{-r_{n+1} t}$. A formal statement of the result can be found as Lemma 4.5.

We now move on to identify and classify the phases of our model.

1.6. Phase behaviour. The flow X is time homogeneous, since the underlying exponential clocks are time homogeneous. From now on we will work only with $(X_{0,t})_{t \geq 0}$. We think of the partition of K into singletons as the initial state of a coalescent, and the flow $t \mapsto X_{0,t}(x)$ as the coalescent mechanism which coagulates the blocks. That is, $x, y \in K$ are in the same block at time t if $X_{0,t}(x) = X_{0,t}(y)$.

DEFINITION 1.9. Let \mathcal{D}_t be the set of $x \in K$ such that, for all w satisfying $x \in K_w$, the clock associated to K_w did not ring during $(0, t]$

Note that $X_{0,t}(x) = x$ for all $x \in \mathcal{D}_t$. Define an equivalence relation on $K \setminus \mathcal{D}_t$ as follows: $x \stackrel{t}{\sim} y$ if and only if the particles which, at time 0 are at x and y respectively, are mapped to the same point of K at time t by the flow defined in Definition 1.8 (i.e. if $X_{0,t}(x) = X_{0,t}(y)$).

DEFINITION 1.10. Let A_t be the set of equivalence classes of $\stackrel{t}{\sim}$.

We call \mathcal{D}_t the *dust* which is present at time t , and for each $x \in \mathcal{D}_t$ we say $\{x\}$ is a *trivial block* at time t . Each element of A_t is a *non-trivial block*, and A_t is the set of non-trivial blocks. The division of K into dust and non-trivial blocks, in the same realisation as Figure 1, is shown in Figure 2.

In Section 2.4, by which time we will have our model defined in full generality, we give a mathematically proper definition of the dust and the (non)-trivial blocks. Until then, we ask the reader to believe that both are well defined and $K = \mathcal{D}_t \uplus (\bigcup A_t)$; every point $x \in K$ is either part of the dust or contained in some non-trivial block.

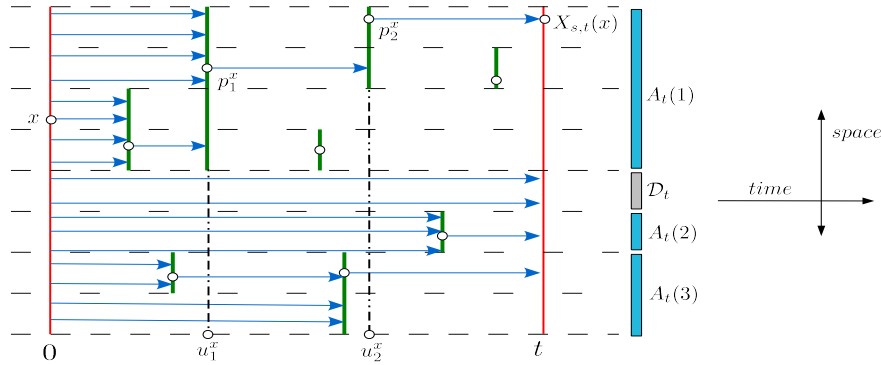


FIG 2. The division of K into dust and non-trivial blocks over time $[0, t]$ is shown. The dust is labelled \mathcal{D}_t , while the non trivial blocks (three, in this realisation) are labelled $A_t = \{A_t(1), A_t(2), A_t(3)\}$. The sequences (u_m^x) and (p_m^x) corresponding to the point $x \in K$ are also shown (to illustrate Definition 2.8). In this example, $N_1^x = 1$, $N_2^x = 2$, and $N_3^x = \infty$.

Loosely speaking, phase transitions occur as we increase the reproduction rate from a sequence where $r_n \downarrow 0$, through to a sequence where $r_n \uparrow \infty$. Informally, a phase transition means a change in qualitative behaviour of the trivial or non-trivial blocks. They may change behaviour together, or one behaviour may change while the other does not.

To be precise, we classify our phases according to whether the total number $|A_t|$ of non-trivial blocks is finite or countably infinite, and whether \mathcal{D}_t is empty, non-empty and null, or takes up a positive fraction of the population. When considering the measure of $\mathcal{D}_t \subseteq K$, we mean with respect to the uniform probability measure on K .

Note that A_t cannot be uncountable since the total number of clocks is countable, and each clock runs at finite rate. When \mathcal{D}_t takes up a positive fraction of the population we say \mathcal{D}_t is *positive*.

A formal definition of the phases can be found in Section 2.4. For now, we define the five phases by means of the following table. A crossed box means that, in the corresponding phase, the corresponding behaviour occurs with positive probability. Table 1.6 (coupled with the note immediately below about the critical phase) details all the behaviour which occurs.

In the critical phase the model has the surprising property that for some *deterministic* $t_0 \in (0, \infty)$,

$$\mathbb{P}[\mathcal{D}_t \text{ is null but non-empty, } \mathcal{A}_t \text{ is countably infinite}] > 0 \text{ for } t < t_0$$

\mathcal{D}_t is:	empty	positive	positive	non-empty, null
A_t is:	finite	finite	infinite	infinite
LOWER SUBCRITICAL	X	X		
UPPER SUBCRITICAL	X		X	
SEMICRITICAL	X			X
CRITICAL	X			X (before t_0)
SUPERCritical	X			

TABLE 1

The phases of our model, and the behaviours occurring in each phase.

but

$$\mathbb{P}[\mathcal{D}_t = \emptyset, \mathcal{A}_t \text{ is finite}] = 1 \text{ for } t > t_0.$$

The time $t_0 \in (0, \infty)$ at which this transition occurs is called the *critical time*.

The critical phase of the model contains (but is not limited to) the case $r_n = c \in (0, \infty)$, where the GWVEs are in fact classical Galton-Watson processes.

We have already argued that the quantities $\sum_n |S|^n r_n$ and $\sum_n r_n$ have a tangible effect on the flow X . A third quantity, which we have not discussed previously, is of equal importance: $\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_1^n r_j$.

The classification of the phases of X is as follows.

THEOREM 2. *Dependent only upon $|S|$ and (r_n) , our model is in precisely one of the five phases. In fact, X is*

- *lower subcritical if and only if $\sum_n |S|^n r_n < \infty$.*
- *upper subcritical if and only if $\sum_n |S|^n r_n = \infty$ and $\sum r_n < \infty$.*
- *semicritical if and only if $\sum r_n = \infty$ and $\limsup_n \frac{1}{n} \sum_1^n r_j = 0$.*
- *critical if and only if $\limsup_n \frac{1}{n} \sum_1^n r_j \in (0, \infty)$*
- *supercritical if and only if $\limsup_n \frac{1}{n} \sum_1^n r_j = \infty$.*

COROLLARY 3. *X comes down from infinity at $t > 0$ if and only if either X is supercritical or X is critical and $t > t_0$.*

As the phase of our model changes, the behaviour of the dust is as expected, in that increasing the intensity of reproduction events reduces the fraction of dust. The lack of monotonicity in the behaviour of the non-trivial blocks is explained as follows. In the lower subcritical phase there are simply not enough events to make anything more than finitely many non-trivial blocks. Then, as the rate increases, there is an intermediate period where we see a countably infinity of non-trivial blocks. Eventually there are so many

reproduction events that they frequently overlap, and we need (a.s.) only finitely many of them to cover K .

In Section 1.4 we commented that when reproduction events are occurring at a high rate, it is common for a larger reproduction event to overwrite the effect of some of the preceding smaller ones. This is borne out by appearance of the lim sup in the formula $\limsup_n \frac{1}{n} \sum_1^n r_j$; from Theorem 2 we see that, when $\sum r_n = \infty$, only the n -level reproduction events for which r_n is large enough to contribute to the lim sup take part in determining the phase.

Our proof of Theorem 2 is outlined in Section 1.10. As we will see, the quantity $\limsup_n \frac{1}{n} \sum_1^n r_j$ characterises the behaviour of the associated GWVEs.

1.7. *Comparison between our model and the Λ -coalescent.* If $\Lambda(\{1\}) > 0$, then the effect on the Λ -coalescent of the atom at 1 is as follows; independently of all other mergers, and at rate $\Lambda(\{1\})$, the Λ -coalescent sees a merger which coagulates the whole population into a single block. Thus, the atom at 1 serves only to obfuscate the behaviour of the Λ -coalescent, and is typically removed. In contrast, in our model, for any $n \in \mathbb{N}$ with $r_n > 0$, there is positive probability of a finite number of level n reproduction events covering K before time t , and causing $|X_{0,t}(K)| \leq |S|^n < \infty$.

In view of the above paragraph, suppose until further notice that $\Lambda(\{1\}) = 0$. As we mentioned in Section 1.1, it is shown in Pitman [1999] that, if the Λ -coalescent comes down from infinity, it does so immediately after time 0. Hence the Λ -coalescent has no equivalent of our critical phase.

Let $\mu^n = \int_0^1 x^n \Lambda(dx)$. Pitman [1999] showed that if $\mu^{-1} = \infty$, the Λ -coalescent has no dust, and if $\mu^{-1} < \infty$, then a non-trivial proportion of the population of the Λ -coalescent is contained within the dust. Hence, the Λ -coalescent has no equivalent of our semicritical phase.

Consider first the case $\mu^{-1} = \infty$. If the Λ -coalescent does not come down from infinity (e.g. the Bolthausen-Sznitman coalescent), then the Λ -coalescent has empty dust and a countable infinity of non-trivial blocks. This behaviour does not occur in our model. Alternatively, if the Λ -coalescent does come down from infinity, then it has empty dust and finitely many atoms, corresponding to our supercritical phase.

Now consider the case $\mu^{-1} < \infty$. It is shown in Freeman [2011] that, if $\mu^{-2} = \infty$, then the Λ -coalescent has a countably infinity of non-trivial blocks, and a positive fraction of the population contained within the dust. Similarly, if $\mu^{-2} < \infty$, then it is well known (e.g. Example 19 in Pitman [1999]) that the Λ -coalescent has only finitely many non-trivial blocks, and has a non-null proportion of the population contained in the dust. Thus the

Λ -coalescent has equivalents of both our upper and lower subcritical phases.

To summarise the previous few paragraphs, the behaviour seen by our model is that of the Λ -coalescent, with following modifications.

1. Playing the role of the cases where $\Lambda(\{1\}) > 0$, we have the (always positive) probability of having only finitely many non-trivial blocks and no dust.
2. There is no possibility in our model of having a countable infinity of non-trivial blocks and empty dust. This behaviour is replaced by our semicritical phase, in which we see a countably infinity of non-trivial blocks and non-empty null dust.
3. The critical phase appears in between the semi- and supercritical phases.

1.8. *The critical phase.* We now turn our attention to the critical phase of our model. By Corollary 3, if X is critical then X comes down from infinity after $t_0 \in (0, \infty)$, but not before. Up until the critical time, there is positive probability of seeing null non-empty dust and an infinity of non-trivial blocks. After the critical time, with probability one, we see only a finite number of non-trivial blocks, and no dust.

COROLLARY 4. *If X is critical then the critical time is given by*

$$t_0 = \frac{\log |S|}{\limsup_n \frac{1}{n} \sum_1^n r_j}.$$

Note that setting $\limsup_n \frac{1}{n} \sum_1^n r_j = 0, \infty$ gives $t_0 = \infty, 0$ respectively, which matches the behaviour seen above and below the critical phase.

Theorem 2 makes no reference to the behaviour of the dust/non-trivial blocks at the critical time t_0 . It turns out that X always comes down from infinity at the critical time, and at t_0 there is no dust and finitely many non-trivial blocks.

COROLLARY 5. *Suppose X is critical. Then, almost surely, $\mathcal{D}_{t_0} = \emptyset$ and A_{t_0} is finite.*

Corollary 5 raises the issue of precisely how the dust disappears in the run up to time t_0 . We give a partial answer to this question in the following section.

1.9. *Fractal dust.* It is natural to ask further questions about the non-empty null dust in the semicritical and critical phases. In these cases \mathcal{D}_t is a

random fractal, in fact it belongs to the large class of random fractals which are, in some sense, stochastic generalizations of iterated function systems.

Recall that an iterated function system (IFS) is a non-empty compact metric space (M, m) equipped with a finite family $(F_i)_{i=1}^n$ of bi-Lipshitz contractions $F_i : M \rightarrow M$. A well known theorem of [Hutchinson \[1981\]](#) states that there is a unique non-empty compact subset \mathcal{A} of M , called the attractor, such that $\mathcal{A} = \bigcup_i F_i(\mathcal{A})$.

Of course, the Cantor set K is one example of an IFS. The set \mathcal{D}_t is, quite simply, K but with (randomly chosen) complexes removed.

IFSs have been generalised in many directions, both deterministically and stochastically, and formulas for the Hausdorff dimension of the corresponding attractors have been obtained in increasing generality. There is a large literature which we will not describe here, and instead refer the reader to [Durand \[2009\]](#), [Mörsters \[2010\]](#) and the references therein. An excellent introduction to fractal geometry can be found in [Falconer \[2003\]](#).

Generality sufficient to cope with \mathcal{D}_t , at least in terms of Hausdorff dimension, seems to have been reached only recently, and in Section 5 we rely heavily on the results of [Durand \[2009\]](#). Results concerning the Hausdorff measure of the attractors are rarer, and a result corresponding to our context does not seem to be known.

In Section 5, to achieve compatibility with [Durand \[2009\]](#), we will impose some extra topological conditions on K . These conditions are satisfied when K is the $|S|$ -part Cantor set, and the result is as follows. A formal statement, with explicit formulae, appears as Theorem 5.2.

Let $\dim_{\mathcal{H}}(A)$ denote the Hausdorff dimension of $A \subseteq \mathbb{R}^d$.

- If X is critical then, conditional on $\mathcal{D}_t \neq \emptyset$, $t \mapsto \dim_{\mathcal{H}}(\mathcal{D}_t)$ decreases linearly over $(0, t_0)$, from the initial value $\dim_{\mathcal{H}}(K)$ at time $0+$ down to the value 0 at t_0- .
- If X is semicritical, or (lower or upper) subcritical, then, conditional on $\mathcal{D}_t \neq \emptyset$, $\dim_{\mathcal{H}}(\mathcal{D}_t) = \dim_{\mathcal{H}}(K)$ for all $t > 0$.

When we say ‘conditional on $\mathcal{D}_t \neq \emptyset$ ’, we mean to condition at fixed time t on the event $\{\mathcal{D}_t \neq \emptyset\}$. Note that, by Theorem 2 and Corollary 5, \mathcal{D}_t is a.s. empty if X is supercritical, or if X is critical and $t \geq t_0$.

The linear decrease of $\dim_{\mathcal{H}}(\mathcal{D}_t)$ over $(0, t_0)$ which occurs in critical phase shows that, as well as becoming empty with increasing probability, the dust also disappears in a geometric sense as $t \uparrow t_0$.

Perhaps the most striking consequence of the above result is that, in the semicritical phase, the action of the reproduction events is sufficiently weak that it does not change the dimension of \mathcal{D}_t . Clearly \mathcal{D}_t , which is a $\dim_{\mathcal{H}}(K)$ -

null set in this phase, must become smaller (in some sense) as time passes, but a description of the way in which this occurs could be very delicate.

1.10. *Outline of the phase behaviour proofs.* Our proof of Theorem 2 comes via several lines of enquiry. Firstly Fubini's theorem produces some useful information, through much the same method as we used in (1.3). Secondly, the spatial structure of K eliminates some of the potential possibilities. For example, Lemma 4.2 shows that, if only finitely many clocks ring, then the dust is either empty or contains a subcomplex K_w . These two approaches are (almost) enough to separate out the lower and upper subcritical phases, using the quantities $\sum_n |S|^n r_n$ and $\sum_n r_n$.

A more powerful tool is required to identify the semicritical, critical and supercritical phases. This is the connection to GWVEs described in Section 1.5. The GWVEs provide information about the behaviour of the dust, in fact, Lemma 4.6 says that (a.s.)

$$(1.5) \quad \mathcal{D}_t = \bigcap_{n=0}^{\infty} \bigcup_{w \in \mathcal{B}_n^t} K_w.$$

Our definition of \mathcal{B}_n^t already suggests that the relationship (1.5) holds. To say that $x \in \bigcap_n \bigcup_{w \in \mathcal{B}_n^t} K_w$ is, by definition of \mathcal{B}^t , precisely the statement that $K_v \ni x$ implies that the clock associated to K_v did not ring during $(0, t]$.

It is clear from (1.5) that the behaviour of $B_n^t = |\mathcal{B}_n^t|$ as $n \rightarrow \infty$ is closely connected to the behaviour of \mathcal{D}_t . In Lemma 4.8 we will show that, with probability one,

$$(1.6) \quad \mathcal{D}_t = \emptyset \text{ if and only if there exists } n \in \mathbb{N} \text{ such that } B_n^t = 0.$$

A GWVE (B_n) is said to be *degenerate* if $\mathbb{P}[\exists n \in \mathbb{N}, B_n = 0] = 1$. Conditions equivalent to degeneracy of B^t were established (for our special case) in Jirina [1976], and are stated here in Lemma 4.10. The most important quantity involved in determining the potential degeneracy of a GWVE is the limiting behaviour of its expectation. By analogy with the classical Galton-Watson process, one might assume that the limiting behaviour of $\mathbb{E}[B_n^t]$ was enough to determine degeneracy, but for general GWVEs this is *not* the case.

Fortunately, Lemma 4.11 shows that the extra complications occur only in the critical phase, at the critical time. In other words, it turns out that degeneracy of B_n^t is equivalent to $\inf_n \mathbb{E}[B_n^t] = 0$, *except* in the critical phase when $t = t_0$. Therefore, for now let us restrict ourselves to examining $\mathbb{E}[B_n^t]$.

Let $m_n^t = \exp(n \log |S| - t \sum_1^n r_j)$ and note that $m_n^t > 0$. The usual calculation for the expected value of a branching process (see, for example, [Fearn 1972](#)) gives us

$$(1.7) \quad \mathbb{E}[B_n^t] = e^{-r_0 t} \prod_{j=1}^n |S| e^{-r_j t} = \mathbb{E}[B_0^t] m_n^t.$$

Since $r_0 \in [0, \infty)$, $\mathbb{E}[B_0^t] = |S| e^{-r_0 t} \in (0, 1]$. Note that $t \mapsto m_n^t$ decreases as t increases, and hence there exists some $t^* \in [0, \infty]$ such that $\inf m_n^t = 0$ if and only if $t > t^*$.

Lemma [4.12](#) shows that the behaviour of $\inf m_n^t$ is governed by the behaviour of $\limsup_n \frac{1}{n} \sum_1^n r_j$. The proof of Lemma [4.12](#) rests on the innocuous rearrangement $m_n^t = \left(\exp \left(\log |S| - t \frac{1}{n} \sum_1^n r_j \right) \right)^n$. Note that this formula hints at the connection to $\limsup_n \frac{1}{n} \sum_1^n r_j$, and also towards the statement of Corollary [4](#). Using Lemma [4.12](#) to restate relevant parts of Theorem [2](#), the connection between B_n^t and the phases is the following. Our model is

- semicritical if and only if $\sum r_n = \infty$ and, for all $t > 0$, $\inf_n m_n^t > 0$.
- critical if and only if there exists $t_0 \in (0, \infty)$ such that
 - for all $t \in (0, t_0)$, $\inf_n m_n^t > 0$.
 - for all $t \in (t_0, \infty)$, $\inf_n m_n^t = 0$.
- supercritical if and only for all $t > 0$, $\inf_n m_n^t = 0$.

The description of the connection between B_n^t and the phases is completed by Corollary [5](#), which deals with the critical time t_0 .

A further piece of information, Lemma [4.9](#), comes out of the fact that, if $\sum r_n < \infty$, then B_n^t grows exponentially as $n \rightarrow \infty$. Using a result of [Biggins and D’Souza \[1992\]](#), we show this fast growth implies that if $\mathcal{D} \neq \emptyset$ then \mathcal{D}_t is non-null.

Although our proof of Theorem [3](#) is essentially subsumed into the proof of the more general Theorem [2](#), it is worth mentioning that the proof of Theorem [3](#) is similar, in a broad sense, to that of its equivalent in [Bertoin and Le Gall \[2006\]](#). In both cases the coalescent is connected to a branching process, and in both cases coming down from infinity corresponds to degeneracy of the branching process. However, as we have seen, our connection to GWVEs is very different to the connection between Λ -coalescents and CSBPs.

1.11. *Layout of Sections 2-5.* The remainder of this article is devoted to a formal definition of the model, and proof of the results given in Section [1](#). Before we proceed further let us make an important note. Recall that we

gave statements of our main results in Sections 1.4-1.8, in the special case where K is the Cantor set.

The statements of Theorem/Corollary 1-5, which appeared in Sections 1.4-1.8, continue to make sense when K assumes the general form defined in Section 2.1. We will not restate these results.

In Section 2.1 we define our geographical space in full generality, and we carry out a proper construction of the model in Section 2.3. Formal definitions of the dust/non-trivial blocks, and of the phases can be found in Section 2.4.

The proofs related to existence of the model appear in Section 3. Proof of Theorem 1 is given in Section 3.3.

After reading Section 2, it is entirely possible to skip the technical proofs of Section 3, and move on to either one of Sections 4 and 5.

The arguments leading up to the phase behaviour proofs can be found in Sections 4.1-4.3. Proof of Theorem 2 and its corollaries can be found in Section 4.4. A formal statement of the discussion in Section 1.9 (on the Hausdorff dimension of \mathcal{D}_t) is given as Theorem 5.2.

2. Definitions. In this section we give a rigorous construction of the full version of our model, and give correspondingly rigorous definitions of the dust, non-trivial blocks, and the five phases.

2.1. Segregated spaces. Let S be a finite set with at least two elements. Let W_n be set of words $w = w_1 w_2 \dots w_n$ of length n , with letters from S (strictly speaking, W_n is the set of ordered sets $\{w_1, \dots, w_n\}$, where $w_i \in S$). We set $W_0 = \{\emptyset\}$, that is we consider the empty word as a word of length 0. Let $W_* = \bigcup_{n=0}^{\infty} W_n$. For each $w = w_1 w_2 \dots w_n \in W_*$ we define $|w| = n$. As usual, if $w = w_1 \dots w_n$ and $i \in S$ then $wi = w_1 \dots w_n i \in W_{n+1}$.

We define the following (non-standard) structure, which is the general form of the geographical space in our model.

DEFINITION 2.1. Let (K, D_K) be a complete metric space, equipped with a family of measurable subsets $(K_w)_{w \in W_*}$ and a probability measure λ . We say K is a segregated space if it satisfies:

- ($\mathcal{K}1$) $K = K_{\emptyset}$ and for all $w \in W_*$, $K_w = \biguplus_{i \in S} K_{wi}$.
- ($\mathcal{K}2$) There exists a sequence $(L_n) \subseteq (0, \infty)$ such that $L_n \rightarrow 0$ and

$$\sup\{D_K(x, y); x, y \in K_w\} \leq L_{|w|}.$$

- ($\mathcal{K}3$) If $|w| = |w'|$ then $\lambda(K_w) = \lambda(K_{w'})$.

Recall that the symbol \uplus denotes disjoint union. Measurability in K is with respect to its Borel σ -field.

In the same spirit as (1.4), ($\mathcal{K}1$) implies that for all $n \in \mathbb{N}$,

$$(2.1) \quad K = \biguplus_{w \in W_n} K_w.$$

We will use ($\mathcal{K}1$) so frequently that it would be impractical to reference it on every application. However, we will not use the other conditions without explicitly saying so. To reconcile with our previous definition, for a general segregated space, K_w is said to be a complex of K with level $|w|$. We extend the definition of subcomplex in the obvious manner.

The point of ($\mathcal{K}2$) is as follows. Suppose (w_n) is a sequence in W_* such that $|w_n| \rightarrow \infty$ and $K_{w_{n+1}} \subseteq K_{w_n}$, and suppose (x_n) is a sequence in K such that $x_n \in K_{w_n}$. Then ($\mathcal{K}2$) implies that (x_n) is Cauchy, and we can use the completeness of K to infer existence of a limit $x_n \rightarrow x \in K$.

Due to ($\mathcal{K}3$), it is natural to think of λ as a uniform measure on K . The measure λ plays no part in the construction of the flow X , but it carries the distinction between the dust and the non-trivial blocks, as the following result shows.

LEMMA 2.2. *For all $w \in W_*$, $\lambda(K_w) > 0$. For all $x \in K$, $\lambda(\{x\}) = 0$.*

PROOF. Let $w \in W_n$. By (2.1), for all $n \in \mathbb{N}$, $K = \biguplus_{v \in W_n} K_v$ and by ($\mathcal{K}3$), $\lambda(K) = \sum_{v \in W_n} \lambda(K_v) = |S|^n \lambda(K_w)$. Since $\lambda(K) = 1$, $\lambda(K_w) > 0$.

Suppose that, for some $x \in K$, $\lambda(\{x\}) = \delta > 0$. Then for all $n \in \mathbb{N}$, $x \in K_w$ for some $w \in W_n$. Hence, by the same calculation as above, $\lambda(K) \geq |S|^n \delta$. Since $n \in \mathbb{N}$ was arbitrary this is a contradiction, so $\lambda(\{x\}) = 0$. ■

EXAMPLE 2.3. *The $|S|$ -part Cantor set, as defined in the introduction, with its Bernoulli measure, is an example of a segregated space. In general, the attractor of an iterated function system which satisfies $\mathcal{A} = \biguplus_i F_i(\mathcal{A})$, equipped with the corresponding Bernoulli measure and with $K_{i_1 \dots i_n} = F_{i_1} \circ \dots \circ F_{i_n}(\mathcal{A})$, is a completely segregated space.*

EXAMPLE 2.4. *Many examples of segregated spaces can be constructed from iterated function systems where the n -level complexes overlap slightly, by arbitrarily choosing which subcomplexes any overlap should belong to. For example, (take λ to be Lebesgue measure on $[0, 1]$ and) note*

$$[0, 1] = [0, 1/2] \cup [1/2, 1] = [0, 1/4] \cup [1/4, 1/2] \cup [1/2, 3/4] \cup [3/4, 1] = \dots$$

We could decide that whenever we see an overlap between complexes the point belongs to the complex which is closer to the origin. Thus we would have $[0, 1]$ decomposed into a disjoint subcomplexes as follows:

$$[0, 1] = [0, 1/2] \uplus (1/2, 1] = [0, 1/4] \uplus (1/4, 1/2] \uplus (1/2, 3/4] \uplus (3/4, 1] = \dots$$

In Example 2.4 the subcomplexes of K may touch (that is, we can have $\overline{K_w} \cap \overline{K_{w'}} \neq \emptyset$ and $K_w \cap K_{w'} = \emptyset$), but in Example 2.3 they cannot. We introduce some terminology to capture the difference. Define

$$\mathcal{O} = \{x \in K; \exists n \in \mathbb{N}, w, w' \in W_n, \text{ such that } w \neq w' \text{ and } x \in \overline{K_w} \cap \overline{K_{w'}}\}.$$

DEFINITION 2.5. *If K satisfies $(\mathcal{K}1)$ – $(\mathcal{K}3)$, and $\mathcal{O} = \emptyset$, then we say K is completely segregated.*

Note that Example 2.3 is completely segregated, whereas Example 2.4 is not. None of our results will require that K is completely segregated, but it is a useful concept and we will make reference to it several times in the sequel.

LEMMA 2.6. *If K is completely segregated then for all $w \in W_*$, K_w is both closed and open.*

PROOF. If $\mathcal{O} = \emptyset$ then for all $n \in \mathbb{N}$, $K = \uplus_{w \in W_n} K_w$ implies that $K = \bigcup_{w \in W_n} \overline{K_w}$, but $\overline{K_w} \cap \overline{K_{w'}} = \emptyset$ for $w \neq w' \in W_n$. Hence $K_w = \overline{K_w}$ is closed. Hence, $K_w = K \setminus (\bigcup_{w' \in W_n \setminus \{w\}} K_{w'})$ is open. ■

Examples like the Cantor set and $[0, 1]$ might give the impression that the dimension of the geographical space plays a role in the genealogy. Our definition of a segregated space does not stipulate anything about the spatial proximity of the complexes. The distance between complexes is encoded in the metric D_K , which affects the dimension of K . However, it is easily seen that the genealogy of the flow X is unaffected when D_K is replaced by any topologically equivalent metric.

That said, if K is isometrically embedded into some ‘standard’ metric space \mathcal{U} (\mathbb{R}^d , for example), then D_K is canonically specified by \mathcal{U} and dimension becomes a useful quantity. It is then possible to analyse the size of the dust \mathcal{D}_t in terms of the change in the dimension of \mathcal{D}_t . We will do precisely this, using the Hausdorff dimension with respect to the Euclidean metric on \mathbb{R}^d , in Section 5.

REMARK 2.7. *This is not the first example of a population model, or even a Fleming-Viot-like process, which is set in a spatial but intrinsically dimensionless context; models on the hierarchical group are well studied. See Dawson and Greven [1999] and the references therein.*

2.2. *Parameters of the model.* For the remainder the article we assume that K is a segregated space. In addition to $|S|$, the parameters in our model are

- A sequence $(r_n)_{n=0}^\infty$ such that $r_n \in [0, \infty)$ and $\exists n, r_n > 0$.
- A non-atomic measure \mathcal{U} on K .

We require the following two conditions on \mathcal{U} .

($\mathcal{K}4$) $\mathcal{U}(K_w) > 0$ for all $w \in W_*$.

($\mathcal{K}5$) *There exists a constant $\alpha \in (0, 1)$ such that for all $w \in W_*$,*

$$\mathcal{U}\left(\bigcup\{K_{wi}; i \in S, \overline{K_{wi}} \not\subseteq K_w\}\right) \leq \alpha \mathcal{U}(K_w).$$

Both these conditions are satisfied by Examples 2.3 and 2.4 with $\mathcal{U} = \lambda$. Before we discuss ($\mathcal{K}4$) and ($\mathcal{K}5$), let us define the reproduction mechanism. Define a measure \mathcal{P} on $W_* \times K$ by, for each $w \in W_*$,

$$(2.2) \quad \mathcal{P}(\{w\} \times A) = r_{|w|} \times \frac{\mathcal{U}(A \cap K_w)}{\mathcal{U}(K_w)}$$

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space equipped with a Poisson point process Π in $\mathbb{R} \times W_* \times K$, of intensity $dt \otimes \mathcal{P}(dw, dy)$ where dt denotes Lebesgue measure. For (measurable) $I \subseteq \mathbb{R}$, $V \subseteq W_*$ and $A \subseteq K$ define

$$\Pi_I = \{(t, w, x) \in \Pi; t \in I\}, \quad \Pi_{I \times V} = \{(t, w, x) \in \Pi; t \in I, w \in V\}.$$

The Poisson point process Π is the mechanism which we informally referred to as exponential clocks, in Section 1.4.

The sequence (r_n) controls the rate at which reproduction events occur, and the measure \mathcal{U} controls the location of the parent. By (2.2), if the reproduction event takes place in K_w , then the parent location is sampled according to the conditional measure of \mathcal{U} on K_w . The condition ($\mathcal{K}4$) is required for our definition of Π , but the point is that the parent sampling has no effect on the rate at which reproduction events occur. Thus \mathcal{U} plays no part in determining the phase of our model.

Loosely speaking, ($\mathcal{K}5$) is satisfied when the points sampled from \mathcal{U} do not concentrate near a non-closed (in the topological sense) edge of any complex. By Lemma 2.6, ($\mathcal{K}5$) is trivially satisfied if K is completely segregated. Condition ($\mathcal{K}5$) prevents pathological examples of the parent sampling mechanism; the need for ($\mathcal{K}5$) will become apparent in Section 2.3.

2.3. *Formal definition of the model.* In this section we give a formal definition of our model, namely the flow X . We begin with some discussion, which motivates the formal definition. Let us briefly suppose that the K is completely segregated and that total event rate is finite; $\sum_n |S|^n r_n < \infty$. Then $\Pi_{[0,t]}$ is almost surely finite and we can represent the process over $[0, t]$ as in Figure 1. By following the arrows in Figure 1, we see precisely where we would like the flow X to map each $x \in K$ to, at time $t > 0$. We will shortly introduce the proper notation for doing this.

What is important to note in Figure 1 is which events affected the final position of the lineages. Consider an event (s, w, y) in a complex K_w of level $|w| = n$ at time $s \in (0, t)$. The event had no effect on the final position (at time t) of any of the lineages if:

- There was an event (s', w', y') such that $s < s' < t$ and $K_w \subseteq K_{w'}$.
- Or, the final event (s', w', y') such that $0 < s' < s$ and $K_w \subset K_{w'}$ had $y' \notin K_w$.

Hence, to work out where a point $x \in K$ should be mapped to (i.e. where, from time s , the individual at x has its ancestor at time $t > s$), we need only consider the following sequence of events.

First, we look for the final level 1 event during $(0, t]$ which affected the point x . If we find one, say (u_1, w_1, p_1) , we then look for the final level 2 event which was after time u_1 , and affected p_1 . And so on. If at any point we don't find a level n event, we simply move up to the next level $n + 1$ and look there. Of course, if $\sum r_n < \infty$ then, at stage m of the induction, p_m might not have been affected by a single event over $(s_m, t]$, in which case we record this and stop the induction.

The formal embodiment of the preceding paragraph is as follows, and a graphical demonstration of the definition can be found in Figure 2.

DEFINITION 2.8. *Fix (x, s, t) with $x \in K$ and $s < t$. Define $(u_0, w_0, p_0) = (s, \emptyset, x)$, and then for as long as $N_{m+1} < \infty$ define the (possibly finite) sequence*

$$\begin{aligned} N_1 &= \inf\{n \geq 0; \exists(u, w, p) \in \Pi_{(s,t] \times W_n} \text{ such that } p_0 \in K_w\} \\ E_1 &= (u_1, w_1, p_1) \\ &\text{where } u_1 = \max\{u \in (s, t]; (s, w, p) \in \Pi_{(s,t] \times W_{N_1}} \text{ and } p_0 \in K_w\}. \end{aligned}$$

$$\begin{aligned} N_{m+1} &= \inf\{n > N_m; \exists(u, w, p) \in \Pi_{[u_m, t] \times W_n} \text{ such that } p_m \in K_w\} \\ E_{m+1} &= (u_{m+1}, w_{m+1}, p_{m+1}) \end{aligned}$$

where $u_{m+1} = \max\{u \in [u_m, t]; (u, w, p) \in \Pi_{[u_m, t] \times W_{N_{m+1}}}\}$
and $p_m \in K_w$.

Recall that $r_n \in [0, \infty)$ and $\mathcal{U}(K) = 1$. As a consequence, for all $w \in W_*$ and $u \in (s, t]$ there is at most one point in the finite set $\Pi_{[s, t]}$ of the form (u, \cdot, \cdot) . Hence Definition 2.8 make sense, almost surely, for an arbitrary single point $x \in K$.

The (finite or infinite) sequence $(E_m) \subseteq \Pi_{(s, t]}$ are the only events which affected the final position of the lineage started from x over time $(s, t]$. At this point we make the key observation; for any $x \in K$, the sequence $(E_m) = (E_m)_{m \geq 1}$ is well defined *even* if we remove the condition that $\sum |S|^n r_n$ be finite. In general, we have the following lemma.

We use the notation $s_n \uparrow s$ to mean that the sequence $(s_n) \subseteq \mathbb{R}$ is strictly monotone increasing, and $s_n \rightarrow s$.

LEMMA 2.9. *If $\sum r_n < \infty$ then (E_m) is almost surely a finite sequence, $(E_m)_1^M$, and $N_{M+1} = \infty$. If $\sum r_n = \infty$, then, almost surely,*

- *the sequence (E_m) is countably infinite, and $N_m \uparrow \infty$.*
- *$u_m \uparrow t$, $|w_m| \uparrow \infty$ and (p_m) converges.*

PROOF. If $\sum_n r_n < \infty$, note each point $x \in K$ is affected by reproduction events at rate $\sum_n r_n < \infty$. Thus the times in between the (y_m) are bounded below by exponential random variables with parameter $\sum_n r_n$.

Now suppose $\sum_n r_n = \infty$. To prove that (E_m) is infinite, note that for any $p \in K$

$$\{(u, w, p) \in \mathbb{R} \times W_* \times K; |w| \geq m, u \in (u_m, t), p \in K_w\}$$

has $dt \otimes \mathcal{P}(dw, dp)$ measure $(t - u_m) \times \sum_{n \geq m} r_n = \infty$, so the probability that p_m is hit by no events in (s_m, t) is zero.

To prove the remaining statements, suppose that for some $\epsilon > 0$, for all $m \in \mathbb{N}$, $u_m \leq t - \epsilon$. Then, by definition of (E_m) , for all $m \geq M$ there are no points in $\Pi_{(t-\epsilon, t]}$, which is a contradiction. Hence $u_m \uparrow t$. Since $|w_m| = N_m$ and $N_{m+1} \geq 1 + N_m$, it is clear that $|w_m| \rightarrow \infty$. Finally, note that K_{w_m} is a decreasing sequence of sets, and in fact $\sup\{m(y, z); y, z \in K_{w_m}\} \leq L_{|w_m|} \rightarrow 0$. For all $k \geq m$, $p_k \in K_{w_m}$, and it follows from ($\mathcal{K}2$) that (p_m) is Cauchy, hence convergent, in K . ■

NOTATION 2.10 (Continuation of Definition 2.8). *The sequence (E_m) depends on x, s and t , and when we need this distinction (which will be most of the time) we write $E_m^{x, s, t} = (u_m^{x, s, t}, w_m^{x, s, t}, p_m^{x, s, t})$. We write $K_{w_m^{x, s, t}} = K_{w_m}^{x, s, t}$.*

We will use the sequence $(E_m^{x,s,t})$ to define $x \mapsto X_{s,t}(x)$ in (2.3), but there are some technicalities to work around before we can do this.

If K is completely segregated then (by Lemma 2.6) all complexes K_w of K are closed. Therefore, it makes intuitive sense that no amount of reproduction events occurring in complexes $K_{w'} \subseteq K_w$ could move particles in the flow from within K_w into $K \setminus K_w$.

If K is not completely segregated then it might be the case that an infinite sequence (u'_m, w'_m, p'_m) of events, with $K_{w'_m} \subseteq K_w$, could have $\lim p'_m \notin K_w$, because it could be that $\lim p'_m \in \bar{K}_w \setminus K_w$. In this case our construction would run into a serious problem; the flow property $X_{s,v} = X_{t,v} \circ X_{s,t}$ would fail. Thanks to ($\mathcal{K}5$), we are able to prove that $\lim p'_m$ is always in K_w , as the following result shows.

Define $\mathcal{C} \subseteq \Omega$ by

$$\mathcal{C} = \left\{ \lim_{m \rightarrow \infty} p_m^{x,s,t} \in K_{w_n}^{x,s,t} \text{ for all } n \in \mathbb{N} \text{ and } x, s, t \text{ such that } |E^{x,s,t}| = \infty \right\}.$$

LEMMA 2.11. $\mathbb{P}[\mathcal{C}] = 1$.

The proof of Lemma 2.11, which includes our only application of ($\mathcal{K}5$), is given in Section 3.1. Note that, by Lemma 2.6, if K is completely segregated then all the K_w are closed and Lemma 2.11 is trivially true.

We stated above that, for fixed $x \in K$, the sequence $(E_m^{x,s,t})$ from Definition 2.8 made sense, almost surely. There are uncountably many $x \in K$, and thus we must check that, almost surely, Definition 2.8 makes sense for all $s < t$ and $x \in K$. To this end define,

$$\begin{aligned} \mathcal{A}_{n,k} &= \{ \Pi_{[-k,k] \times W_n} \text{ is finite} \} \\ \mathcal{B} &= \{ \text{for all } u \in \mathbb{R}, \Pi_{\{u\}} \text{ is at most a single point} \} \end{aligned}$$

By standard properties of Poisson point processes, $\mathbb{P}[\mathcal{A}_{n,k}] = 1$. For any two $w, w' \in W_*$, the probability of the processes $\Pi_{\mathbb{R} \times \{w\}}$ and $\Pi_{\mathbb{R} \times \{w'\}}$ causing a reproduction event at a common time is 0. Since W_* is countable, in fact also $\mathbb{P}[\mathcal{B}] = 1$. For realisations $\omega \in (\cap_{n,k} \mathcal{A}_{n,k}) \cap \mathcal{B}$, Definition 2.8 makes sense for all x, s, t .

Let $\mathcal{A} = (\cap_{n,k} \mathcal{A}_{n,k}) \cap \mathcal{B} \cap \mathcal{C}$. By Lemma 2.11, $\mathbb{P}[\mathcal{A}] = 1$. Because of Lemma 2.11, it is advantageous to only use $(E_m^{x,s,t})$ for $\omega \in \mathcal{A}$.

Briefly suppressing dependence on (s, t) , note that $(K_{w_m}^x)$ (for $\omega \in \mathcal{A}$) is a decreasing infinite sequence of sets. We have also that $p_m^x \in K_{w_m}^x$ and $\sup\{m(y, z); y, z \in K_{w_m}^x\} \leq L_{|w_m^x|}$. If (p_m^x) is infinite then by completeness of K and ($\mathcal{K}2$), (p_m^x) is convergent.

So, for $\omega \in \mathcal{A}$, define

$$(2.3) \quad X_{s,t}(x) = \begin{cases} x & \text{if } N_1^{x,s,t} = \infty \\ p_M^{x,s,t} & \text{if } N_{M+1}^{x,s,t} = \infty, \text{ for } M \in \mathbb{N} \\ \lim_{m \rightarrow \infty} p_m^{x,s,t} & \text{if } (E_m^{x,s,t}) \text{ is infinite.} \end{cases}$$

and for $\omega \notin \mathcal{A}$ simply set $X_{s,t} = \text{id}$, the identity function on K . The first case of (2.3) corresponds to the dust, and the other two cases correspond to the non-trivial blocks.

Recall that our existence result, Theorem 1, did not specify any continuity/measurability properties of the flow X (although these would normally be present in a modern definition of a stochastic flow, see for example [Le Jan and Raimond 2004](#)). We make some remarks in this direction.

REMARK 2.12 (On particle paths). *In Section 1.3 we gave much attention to the behaviour of ancestral lineages. Note that there is no group structure, or even an intrinsic binary operation, on K , so the concept of a Lévy/compound Poisson process in K is not well defined. We will not make a formal statement of the equivalent result for our model, however:*

For each $x \in K$, $t \mapsto X_{0,t}(x)$ is a càdlàg function, and $\omega \mapsto X_{0,\cdot}(x)$ is a random variable in the space of càdlàg K -valued paths (with the usual weak topology). If \mathcal{J} is the set of discontinuities of $t \mapsto X_{0,t}(x)$, then \mathcal{J} is a.s. locally finite if $\sum r_n < \infty$, and a.s. dense in $(0, \infty)$ if $\sum r_n = \infty$. Proof of these two statements is not much more than a long exercise in manipulating the definitions, and is not included in this article.

REMARK 2.13 (On time continuity). *It holds that for all $s_0 < t_0$,*

$$\lim_{s \rightarrow s_1, t \rightarrow t_1} \mathbb{E} [|\varphi_{s,t}, \varphi_{s_1,t_1}|] = 0.$$

To see this, note that for all $n \in \mathbb{N}$, the rate at which reproduction events occur in complexes K_w such that $|w| \leq n$ is $\sum_0^n |S|^j r_j < \infty$. Hence, in any small time interval, with high probability we only see reproduction events with levels $m \geq n$, where n is large. Such reproduction events do not alter the position of particles within the flow by more than $L_m \leq L_n$. By (K2), $L_n \rightarrow 0$ as $n \rightarrow \infty$.

REMARK 2.14 (On spatial continuity). *For some segregated spaces K , with positive probability, $x \mapsto X_{0,t}(x)$ has discontinuities at deterministic points in space. It is easily seen (as in Figure 1) that Example 2.4 has this behaviour.*

2.4. *Formal definition of the phases.* For each t we partition $K = \mathcal{D}_t \uplus (K \setminus \mathcal{D}_t)$, where, in the notation of Section 2.3, $\mathcal{D}_t = \{x \in K; N_1^{x,0,t} = \infty\}$. Note that $N_1^{x,s,t} = \infty$ if and only if $(E_m^{x,s,t}) = \emptyset$.

The set \mathcal{D}_t is only defined almost surely (i.e. when $\omega \in \mathcal{A}$), which is irritating but has no serious consequences. It does mean that every statement we make about \mathcal{D}_t should be qualified by the phrase ‘almost surely’.

By definition of $N^{x,0,t}$, for all \mathcal{D}_t we have $X_{0,t}(x) = x$. As in Section 1.6, for $x \in \mathcal{D}_t$, we say $\{x\}$ is a trivial block of $X_{0,t}$, and we refer to \mathcal{D}_t as the dust of $X_{0,t}(x)$. By definition of X (and Lemma 2.11) if $x \notin \mathcal{D}_t$ then $K_{w_1}^{x,0,t} = \{y \in K; X_{0,t}(x) = X_{0,t}(y)\}$. It is for the reason that we refer to $K_{w_1}^{x,0,t}$ as a non-trivial block of $X_{0,t}$.

For $t > 0$ define λ_t^* to be the measure on K defined by $\lambda_t^*(\cdot) = \lambda(X_{0,t}^{-1}(\cdot))$, and let

$$\mathcal{A}_t = \{y \in K; y \text{ is an atom of } \lambda_t^*\}.$$

To reconcile this with our notation from Section 1.6, the non-trivial blocks $B \in \mathcal{A}_t$ are precisely the pre-images $X_{0,t}^{-1}(y)$ of $y \in \mathcal{A}_t$. It follows from Lemma 2.16 (below) that $|\mathcal{A}_t| = |\mathcal{A}_t|$.

REMARK 2.15. *As a consequence of Lemma 2.11, our reproduction mechanism does not notice whether subcomplexes of any K_w are touching or have positive distance between them. In other words, in so far as trivial/non-trivial blocks as concerned, the genealogy of X is unaffected by whether K is completely segregated or just segregated.*

From now on it will be convenient for us to write $f^{-1}(\{x\}) = f^{-1}(x)$ for functions f , and $\mu(\{x\}) = \mu(x)$ for measures μ . The measure λ_t^* carries all the information about the action of $X_{0,t}$ on K , as the following result shows.

LEMMA 2.16. *The following statements hold:*

1. For all $x \in \mathcal{D}_t$, $X_{0,t}^{-1}(x) = \{x\}$, and $\lambda_t^*(x) = 0$.
2. $\lambda_t^*(\mathcal{D}_t) = \lambda(\mathcal{D}_t)$.
3. If $y \in \mathcal{A}_t$ then $y \notin \mathcal{D}_t$, and $X_{0,t}^{-1}(y) = K_{w_1}^{x,0,t}$ for some $x \in K$.
4. For all $x \in K \setminus \mathcal{D}_t$, $X_{0,t}(x) \in \mathcal{A}_t$.

PROOF. Let $x \in \mathcal{D}_t$, then by definition $X_{0,t}(x) = x$. Suppose $z \in K \setminus \{x\}$ and $X_{0,t}(x) = z$. Then $(E_m^{z,0,t}) \neq \emptyset$. If $x \notin K_{w_1}^{z,0,t}$ then by Lemma 2.11 we would have $X_{0,t}(x) \neq X_{0,t}(z)$, so we must have $x \in K_{w_1}^{z,0,t}$. But this contradicts $x \in \mathcal{D}_t$. Hence in fact $X_{0,t}^{-1}(x) = x$. Hence $\lambda_t^*(\mathcal{D}_t) = \lambda(\mathcal{D}_t)$, and it follows immediately from Lemma 2.2 that $\lambda_t^*(x) = \lambda(x) = 0$.

If y is an atom of λ_t^* we must have $y = X_{0,t}(x)$ for some $x \in K \setminus \mathcal{D}_t$. By definition of X , we have also that $K_{w_1}^{x,0,t} \subseteq X_{0,t}^{-1}(y)$. Hence $y \notin \mathcal{D}_t$. If $z \in X_{0,t}^{-1}(y)$ and $z \notin K_{w_1}^{x,0,t}$, then Lemma 2.11 implies $X_{0,t}(z) \neq X_{0,t}(x) = y$, so in fact $K_{w_1}^{x,0,t} = X_{0,t}^{-1}(y)$.

Finally, suppose that $x \in K \setminus \mathcal{D}_t$ and let $y = X_{0,t}(x)$. By definition of X and Lemma 2.11 we have $K_{w_1}^{x,0,t} = X_{0,t}^{-1}(y)$, which by Lemma 2.2 implies that $\lambda_t^*(y) > 0$. ■

Define

$$\begin{aligned}\mathcal{P}_1^t &= \{\mathcal{D}_t = \emptyset, \mathcal{A}(\lambda_t^*) \text{ is finite}\} \\ \mathcal{P}_2^t &= \{\lambda_t^*(\mathcal{D}_t) > 0, \mathcal{A}(\lambda_t^*) \text{ is finite}\} \\ \mathcal{P}_3^t &= \{\lambda_t^*(\mathcal{D}_t) > 0, \mathcal{A}(\lambda_t^*) \text{ is countably infinite}\} \\ \mathcal{P}_4^t &= \{\lambda_t^*(\mathcal{D}_t) = 0, \mathcal{D}_t \neq \emptyset, \text{ and } \mathcal{A}(\lambda_t^*) \text{ is countably infinite}\}.\end{aligned}$$

We write $\mathbb{P}\langle A * B \rangle = 1$ to mean that $\mathbb{P}[A \cup B] = 1$ and both $\mathbb{P}[A] > 0$, $\mathbb{P}[B] > 0$. The phases are formally defined as follows.

DEFINITION 2.17. *We say that X is*

- *Lower subcritical if for all $t > 0$, $\mathbb{P}\langle \mathcal{P}_1^t * \mathcal{P}_2^t \rangle = 1$.*
- *Upper subcritical if for all $t > 0$, $\mathbb{P}\langle \mathcal{P}_1^t * \mathcal{P}_3^t \rangle = 1$.*
- *Semicritical if for all $t > 0$, $\mathbb{P}\langle \mathcal{P}_1^t * \mathcal{P}_4^t \rangle = 1$.*
- *Critical if there exists $t_0 > 0$ such that*
 - *for all $t \in (0, t_0)$, $\mathbb{P}\langle \mathcal{P}_1^t * \mathcal{P}_4^t \rangle = 1$.*
 - *for all $t \in (t_0, \infty)$, $\mathbb{P}[\mathcal{P}_1^t] = 1$.*
- *Supercritical if $\forall t > 0$, $\mathbb{P}[\mathcal{P}_1^t] = 1$.*

3. Existence. Recall that $E_m^{x,s,t} = (u_m^{x,s,t}, w_m^{x,s,t}, p_m^{x,s,t})$ was defined by Definition 2.8 (and Notation 2.10). Note that $|w_m^{x,s,t}| = N_m^{x,s,t}$, $p_m^{x,s,t} \in K_{w_{m+1}}^{x,s,t}$, and $u_m^{x,s,t} < u_{m+1}^{x,s,t}$.

For all $\omega \in \mathcal{A}$, $X_{s,t}(x)$ is defined by (2.3). This is an almost sure definition of $(X_{s,t})$, and for the \mathbb{P} -null subset $\Omega \setminus \mathcal{A}$ which (2.3) does not cover, $X_{s,t}$ was defined to be the identity function $\iota : K \rightarrow K$.

3.1. *Proof of Lemma 2.11.* Let

$$H' = \{(x, s, t) \in K \times \mathbb{R}^2; |E^{x,s,t}| = \infty\}$$

and define an equivalence relation on H' by

$$(x, s, t) \sim (x', s', t') \Leftrightarrow E_1^{x,s,t} = E_1^{x',s',t'}.$$

Denote the equivalence class of (x, s, t) under \sim by $[x, s, t]$. Note that for all $(x', s', t') \in [x, s, t]$, $(E_m^{x, s, t}) = (E_m^{x', s', t'})$. We thus write $u_m^{x, s, t} = u_m^{[x, s, t]}$, and similarly for w and p .

Note also that if $x' \in K_{w_1}^{x, s, t}$, $s' < u_1^{x, s, t}$ and $u^{x, s, t} < t' < t$ then $(x', s', t') \in [x, s, t]$. Let $(\hat{x}_k)_{k \in \mathbb{N}}$ be a deterministic countable subset of K , such that that for all $w \in W_*$ there is $k \in \mathbb{N}$ such that $x_k \in K_w$. Let (\hat{s}_k, \hat{t}_k) be a countably dense deterministic subset of $\{(s, t) \in \mathbb{R}^2; s < t\}$. With a slight abuse of notation, enumerate their Cartesian product as

$$(\hat{x}_k, \hat{s}_k, \hat{t}_k)_{k \in \mathbb{N}}.$$

Thus, for all equivalence classes $[x, s, t]$ there exists (random) $k \in \mathbb{N}$ such that $(\hat{x}_k, \hat{s}_k, \hat{t}_k) \in [x, s, t]$.

Hence,

$$\begin{aligned} & \{\exists x, s, t \in H', \exists n \in \mathbb{N}, \lim_{m \rightarrow \infty} p_m^{x, s, t} \notin K_{w_n}^{x, s, t}\} \\ &= \{\exists [\hat{x}_k, \hat{s}_k, \hat{t}_k] \subseteq H', \exists n \in \mathbb{N}, \lim_{m \rightarrow \infty} p_m^{[x, s, t]} \notin K_{w_n}^{[x, s, t]}\} \\ (3.1) \quad &= \bigcup_{k \in \mathbb{N}} \bigcup_{n \in \mathbb{N}} \{\lim_{m \rightarrow \infty} p_m^{[\hat{x}_k, \hat{s}_k, \hat{t}_k]} \notin K_{w_n}^{[\hat{x}_k, \hat{s}_k, \hat{t}_k]}\}. \end{aligned}$$

By $(\mathcal{K}1)$, $\lim_m p_m^{[\hat{x}_k, \hat{s}_k, \hat{t}_k]} \notin K_{w_n}^{[\hat{x}_k, \hat{s}_k, \hat{t}_k]}$ occurs only if, for all $m \in \mathbb{N}$, $p_m^{[\hat{x}_k, \hat{s}_k, \hat{t}_k]}$ is such that

$$\overline{K_{w_{m+1}}^{[\hat{x}_k, \hat{s}_k, \hat{t}_k]}} \setminus K_{w_n} \neq \emptyset.$$

By $(\mathcal{K}5)$, for all $k, n \in \mathbb{N}$, for each fixed m the probability of this occurring is bounded above by $\alpha < 1$. The points $p_m^{[\hat{x}_k, \hat{s}_k, \hat{t}_k]}$ are independent random variables, each sampled according to the conditional measure of \mathcal{U} on $K_{w_m}^{[\hat{x}_k, \hat{s}_k, \hat{t}_k]}$. Hence, $\mathbb{P} \left[\lim_{m \rightarrow \infty} p_m^{[\hat{x}_k, \hat{s}_k, \hat{t}_k]} \notin K_{w_n}^{[\hat{x}_k, \hat{s}_k, \hat{t}_k]} \right] = 0$, and it follows from (3.1) that

$$\mathbb{P} [\exists x, s, t \in H', \exists n \in \mathbb{N}, \lim_{m \rightarrow \infty} p_m^{x, s, t} \notin K_{w_n}^{x, s, t}] = 0.$$

Thus $\mathbb{P}[\mathcal{C}] = 1$. ■

3.2. Measurability. Recall that our underlying Poisson point process Π is defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Throughout this section we denote the dependence on $\omega \in \Omega$ of X by writing $X_{0,t}(\cdot)(\omega)$. Let $\mathbb{B}(K)$ denote the Borel σ -algebra on K , and recall that D_K denotes the metric on K .

LEMMA 3.1. *Let $-\infty < s < t < \infty$. For all $\omega \in \Omega$, $X_{s,t} : K \rightarrow K$ is a measurable function.*

PROOF. Since our model is time homogeneous, it suffices to consider the case $s = 0$. If $\omega \notin \mathcal{A}$ then $X_{0,t}$ is the identity function, which is measurable.

If $\omega \in \mathcal{A}$, then $X_{0,t}|_{\mathcal{D}_t}$ is also the identity function, which is measurable. On $K \setminus \mathcal{D}_t$, $X_{0,t}$ has a countable range, which by Lemma 2.16 is precisely the set \mathcal{A}_t . Also by Lemma 2.16, if $y \in \mathcal{A}_t$ then $X_{0,t}^{-1}(y) = K_w$ for some $w \in W_*$. By definition of a segregated space, every K_w is measurable. Thus $K \setminus \mathcal{D}_t$ is a countable union of measurable sets, and is itself measurable. It follows immediately that \mathcal{D}_t is measurable.

Thus the restriction $X_{0,t}$ is measurable on both \mathcal{D}_t and $K \setminus \mathcal{D}_t$, both of which are measurable sets. It follows that $X_{0,t}$ is measurable. ■

LEMMA 3.2. *The Borel σ -algebra on K is generated by $(K_w)_{w \in W_*}$.*

PROOF. By Definition 2.1, each K_w is measurable, so it is clear that $\sigma(K_w; w \in W_*) \subseteq \mathbb{B}(K)$. We will now prove the reverse inclusion.

Since K is a separable metric space, any open subset of K can be written as a union of only countably many open balls of K . Hence $\mathbb{B}(K)$ is generated by the open balls of K . So the proof is complete if we can show that any open ball of K is contained in $\sigma(K_w; w \in W_*)$.

To this end, let $B_r(x) = \{y \in K; |y - x| < r\}$ be a fixed but arbitrary open ball in K . By ($\mathcal{K}1$), for each $y \in K$ and $n \in \mathbb{N}$, let $K_{(y,n)}$ be the unique complex K_w of K such that $|w| = n$ and $y \in K_w$. Note that

$$(3.2) \quad B_r(x) \supseteq \bigcup \{K_w; w \in W_*, K_w \subseteq B_r(x)\}$$

is tautologically true, and, since W_* is countable, the union on the right is countable. Now, suppose that $y \in B_r(x)$. Since $B_r(x)$ is open, for some $\epsilon > 0$ we have $B_\epsilon(y) \subseteq B_r(x)$. By ($\mathcal{K}2$), for some sufficiently large $n \in \mathbb{N}$ we have $K_{w(y,n)} \subseteq B_\epsilon(y) \subseteq B_r(x)$. However, this implies that $K_{w(y,n)} \in \{K_w; w \in W_*, K_w \subseteq B_r(x)\}$, so as $y \in \bigcup \{K_w; w \in W_*, K_w \subseteq B_r(x)\}$. Hence, in fact (3.2) is an equality, and thus $B_r(x) \in \sigma(K_w; w \in W_*)$. The proof is complete. ■

LEMMA 3.3. *For each $s < t$, $(x, \omega) \mapsto X_{s,t}(x)(\omega)$ is a measurable function from $K \times \Omega \rightarrow K$.*

PROOF. As in the proof of Lemma 3.1, it suffices to consider the case $s = 0$. For $v \in W_*$ let $\mathcal{W}(v) = \{w' \in W_*; K_v \subseteq K_{w'}\}$.

Fix $w \in W_*$. We note

$$\begin{aligned} & \{(x, \omega) \in K \times \Omega; X_{0,t}(x)(\omega) \in K_w\} \\ &= \bigcup_{v \in W_*} K_v \times \{\omega \in \Omega; X_{0,t}(K_v) \subseteq K_w\} \\ &= \bigcup_{v \in W_*} K_v \times \left[\{\omega \in \Omega \setminus \mathcal{A}; K_u \subseteq K_v\} \uplus \left(\mathcal{A} \cap \right. \right. \\ & \quad \left. \left. \{\omega \in \Omega; \nexists (u', w', p') \in \Pi_{(0,t]}, K_v \subseteq K_{w'}, p' \notin K_w, \Pi_{[u',t] \times \mathcal{W}(v)} = \emptyset\} \right) \right] \end{aligned}$$

Note that $\{\omega \in \Omega \setminus \mathcal{A}; K_u \subseteq K_v\}$ is either empty or equal to the measurable set $\Omega \setminus \mathcal{A}$. From the representation above, it follows that $\{(x, \omega) \in K \times \Omega; X_{0,t}(x)(\omega) \in K_w\}$ is an element of the product σ -algebra on $K \times \Omega$. Lemma 3.2 completes the proof. ■

3.3. Proof of Theorem 1. By Lemma 3.1, $X_{s,t} \in \mathbb{M}$. We omit a formal proof of the fact that $\omega \mapsto X_{s,t}(\cdot)(\omega)$ is an \mathbb{M} -valued random variable, since it is clear from the definition that (for fixed s, t) we have only used the \mathcal{F} -measurable set \mathcal{A} , countably many (deterministic) operations, and the random variable Π .

The remainder of the proof comes in several parts, which correspond to the bullet points in the statement of Theorem 1.

PART 1. Let $s < t < v$ and fix $x \in K$. If $\omega \notin \mathcal{A}$ then we trivially have $X_{s,v} = X_{t,v} \circ X_{s,t}$. So, suppose $\omega \in \mathcal{A}$, and for convenience write $y = X_{s,t}(x)$. When necessary we will emphasise the dependence with $y = y^{x,s,t}$. We divide into three cases.

If $N_1^{x,s,t} = N_1^{y,t,v} = \infty$, then for all $x \in M$, $X_{s,t}(x) = x = y$ and $X_{t,v}(y) = y$. Since $x = y$, $N_1^{x,s,v} = \infty$ and $X_{s,t}(x) = X_{t,v}(x) = X_{s,v}(x)$, so $X_{t,v} \circ X_{s,t}(x) = X_{s,v}(x)$.

If $N_1^{x,s,t} = \infty$ and $N_1^{y,t,v} < \infty$, then $X_{s,t}(x) = x = y$ and hence we must have $u_1^{x,s,t} \geq v$. Hence $(E_m^{x,t,v}) = (E_m^{x,s,v})$ and thus $X_{t,v}(X_{s,t}(x)) = X_{t,v}(x) = X_{s,t}(x)$.

If $N_1^{x,s,t} < \infty$, then we have $N_1^{x,s,v} < \infty$. Let

$$\mathcal{C}^{s,t,v} = \{x \in K; \exists m, u_m^{x,s,v} \geq t\}.$$

If $x \notin \mathcal{C}^{s,t,v}$ then $u_m^{x,s,v} < t$ for all m , so from the definitions we have $(E_m^{x,s,t}) = (E_m^{x,s,v})$. Hence $X_{t,v}(x) = X_{s,t}(x) = y$. Suppose it was the case that $(u_1^{y,t,v}, w_1^{y,t,v}, p_1^{y,t,v}) \in (E^{y,t,v})$. Note $y \in K_{w_1}^{x,s,v}$ so we must have $(u_1^{y,t,v}, w_1^{y,t,v}, p_1^{y,t,v}) \in (E_m^{x,s,v})$, which is a contradiction since $u_1^{y,t,v} \geq t$. Hence $(E_m^{y,t,v})$ is empty, and $X_{t,v} = \iota$. Thus, $X_{t,v}(X_{s,t}(x)) = X_{s,v}(x)$.

If $x \in \mathcal{C}^{s,t,v}$, let $M = \max\{m; u_m^{x,s,v} < t\}$ (which is well defined since $(u_m^{x,s,t})$ is strictly increasing), and from the definitions note that $(E_m^{x,s,t})_1^M = (E_m^{x,s,v})_1^M$.

By definition of M we have $u_{M+1}^{x,s,v} \geq t$ and, since $p_M^{x,s,v} = p_M^{x,s,t}$, it holds that $N_{M+1}^{x,s,v} \leq N_{M+1}^{x,s,t}$. Hence $K_{w_{M+1}}^{x,s,t} \subseteq K_{w_{M+1}}^{x,s,v}$. By definition, $p_M^{x,s,t} \in K_{w_{M+1}}^{x,s,t}$ and, we have also that $(K_{w_m}^{x,s,t})$ is decreasing. We have already commented that $K_{w_{M+1}}^{x,s,t} \subseteq K_{w_{M+1}}^{x,s,v}$, so it follows from Lemma 2.11 that $y^{x,s,t} \in K_{w_{M+1}}^{x,s,v}$.

Since both y and $p_M^{x,s,v}$ are elements of $K_{w_{M+1}}^{x,s,v}$, there is no $(u, w, p) \in (E_m^{x,s,v})$ such that $|w| < N_{M+1}^{x,s,v}$ - such a (u, w, p) would also have featured in $(E_m^{x,s,t})$, which contradicts the definition of M . Also, there are no $(u, w, p) \in (E_m^{x,s,v})$ such that $u > u_{M+1}^{x,s,v}$ and $y \in K_w$ - such a (u, w, p) would feature in $(E_m^{x,s,t})$, which contradicts the definition of $u_{M+1}^{x,s,v}$.

Combining the results of previous two sentences, $(u_{M+1}^{x,s,v}, w_{M+1}^{x,s,v}, p_{M+1}^{x,s,v}) = (u_1^{y,t,v}, w_1^{y,t,v}, p_1^{y,t,v})$. Hence $(E_m^{x,s,v})_{m \geq M+1} = (E_k^{x,s,v})_{k \geq 1}$, which implies that $X_{t,v}(y) = X_{s,v}(x)$. This completes the third case.

Since x and ω were arbitrary, in all cases we have that for all $\omega \in \Omega$, $X_{s,v} = X_{t,v} \circ X_{s,t}$.

PART 2: Let $s_1 < t_1 \leq s_2 < t_s$. Since $\Pi_{(s_1,t_1]}$ and $\Pi_{(s_2,t_2]}$ are independent, and the construction of $X_{s,t}$ depended only on $\Pi_{s,t}$, it follows immediately that X_{s_1,t_1} and X_{s_2,t_2} are independent.

PART 3: Let $s_1 < t_1$ and $s_2 < t_2$ with $t_1 - s_1 = t_2 - s_2$. Then $\Pi_{(s_1,t_1]}$ and $\Pi_{(s_2,t_2]} - (t_2 - t_1)$ are identical in law, from which it follows that X_{s_1,t_1} and X_{s_2,t_2} are also identical in law. ■

4. The rate of reproduction. We now work towards proof of Theorem 2 and its corollaries.

4.1. *Dust and atoms.* In this section we give some preliminary results, using Fubini's theorem and the spatial structure of K .

LEMMA 4.1. *For all $t > 0$,*

- *If $\sum r_n = \infty$ then $\mathbb{P}[\lambda_t^*(\mathcal{D}_t) = 0] = 1$.*
- *If $\sum r_n < \infty$ then $\mathbb{P}[\lambda_t^*(\mathcal{D}_t) > 0] > 0$.*

PROOF. Fix $t > 0$. By Fubini's theorem, which applies by Lemma 3.1,

$$\mathbb{E}[\lambda(\mathcal{D}_t)] = \mathbb{E}\left[\int_K \mathbb{1}\{x \in \mathcal{D}_t\} \lambda(dx)\right] = \int_K \mathbb{P}[x \in \mathcal{D}_t] \lambda(dx).$$

If $\sum r_n < \infty$ then for each $x \in K$, $\mathbb{P}[x \in \mathcal{D}_t] > 0$. Hence $\mathbb{P}[\lambda(\mathcal{D}_t) > 0] > 0$. Alternatively, if $\sum r_n = \infty$ we have $\mathbb{P}[x \in \mathcal{D}_t] = 0$, so as $\mathbb{P}[\lambda(\mathcal{D}_t) = 0] = 1$. The stated results follow since by Lemma 2.16 we have $\lambda_t^*(\mathcal{D}_t) = \lambda(\mathcal{D}_t)$. ■

LEMMA 4.2. *Suppose λ_t^* has finitely many atoms. Then, almost surely, either $\mathcal{D}_t = \emptyset$ or, for some $w \in W_*$, $K_w \subseteq \mathcal{D}_t$. In the latter case, $\lambda_t^*(\mathcal{D}_t) > 0$.*

PROOF. Consider $\omega \in \mathcal{A}$. Suppose that λ_t^* has finitely many atoms and $\mathcal{D}_t \neq \emptyset$. Let $A = \{y(i); i = 1, \dots, N\} \subseteq K$ be the set of atoms of λ_t^* . By Lemma 2.16, for each $y(i) \in A$ there is some $x(i) \in K$ such that $X_{0,t}^{-1}(y(i)) = K_{w_1}^{x(i),0,t}$. Let $n = \max \{|w_1^{x(i),0,t}|; i = 1, \dots, N\}$. By $(\mathcal{K}1)$, each $K_{w_1}^{x(i),0,t}$ can be written as the union of finitely many K_w such that $w \in W_n$. Since $\mathcal{D}_t \neq \emptyset$ we must have $\bigcup_i K_{w_1}^{x(i),0,t} \subset K$, and thus there is some $w \in W_n$ such that $K_w \subseteq \mathcal{D}_t$. By Lemma 2.16, $\lambda_t^*(\mathcal{D}_t) = \lambda(\mathcal{D}_t) \geq \lambda(K_w)$. It follows by Lemma 2.2 that $\lambda_t^*(\mathcal{D}_t) > 0$. ■

LEMMA 4.3. *Suppose $\sum |S|^n r_n = \infty$. Then $\mathbb{P}[\exists w \in W_*, K_w \subseteq \mathcal{D}_t] = 0$.*

PROOF. Fix $t > 0$. Since W_* is countable, the lemma follows if we can show that $\mathbb{P}[K_w \subseteq \mathcal{D}_t] = 0$ for an arbitrary $w \in W_*$. So fix $w \in W_*$, and set $n = |w|$. The rate at which K_w is affected by reproduction events is

$$\begin{aligned} \int_{W^* \times K} \mathbb{1}\{K_w \cap K_{w'} \neq \emptyset\} \mathcal{P}(dw', dp) &= \sum_{w' \in W_*} r_{|w|} \mathbb{1}\{K_w \cap K_{w'} \neq \emptyset\} \\ &\geq \sum_{w' \in W_*} r_{|w|} \mathbb{1}\{K_{w'} \subseteq K_w\} \end{aligned}$$

Now, by $(\mathcal{K}1)$, $K_{w'} \subseteq K_w$ if and only if $w' = wv$ for some $v \in W_*$. Hence,

$$\sum_{w' \in W_*} r_{|w|} \mathbb{1}\{K_{w'} \subseteq K_w\} = \sum_{v \in W_*} r_{|wv|} = \frac{1}{|S|^n} \sum_{m=n}^{\infty} |S|^m r_m = \infty$$

It follows immediately that (with probability one) K_w is affected by a reproduction event during $(0, t]$ for any $t > 0$. Hence $\mathbb{P}[K_w \subseteq \mathcal{D}_t] = 0$. ■

LEMMA 4.4. *If $\mathcal{D}_t = \emptyset$ then \mathcal{A}_t is finite.*

PROOF. We will first prove that K cannot be expressed as the disjoint union of a countably infinite number of its subcomplexes.

Suppose that $K = \biguplus_{i \in \mathbb{N}} K_{w_i}$. Since the union is disjoint, the sequence (w_i) does not have any repeated terms. Hence $|w(i)| \rightarrow \infty$ as $i \rightarrow \infty$. Choose some sequence $(x_i) \subseteq K$ such that $x_i \in K_{w_i}$ for all i . A Bolzano-Weierstrauss type argument (using only $(\mathcal{K}1)$ and completeness) shows that (x_i) has a convergent subsequence, say $x_{n_i} \rightarrow y \in K$.

Since $y \in K$, there exists $j \in \mathbb{N}$ such that $y \in K_{w_j}$. By Lemma 2.6 K_{w_j} is open, and hence for some $\epsilon > 0$ the open ball $B_\epsilon(y) \subseteq K_{w_j}$. Let i_1 and i_2 be such that $i_1 < i_2$ and $|x_{n_{i_1}} - y|, |x_{n_{i_2}} - y| < \epsilon$. Hence $x_{n_{i_1}}, x_{n_{i_2}} \in K_{w_j}$, which implies $x_{n_{i_1}} = x_{n_{i_2}}$. But the sequence (x_i) did not have repeated terms, and thus we have a contradiction.

Now, suppose \mathcal{A}_t is infinite, in particular say $\mathcal{A}_t = \{x_i \in K; i \in \mathbb{N}\}$, where $x_i = x_j$ if and only if $i = j$. By Lemma 2.16, for each i we have some $y_i \in K$ such that $K_{w_1}^{y_i, 0, t} = X_{0, t}^{-1}(x_i)$. Thus $X_{0, t}(y_i) = x_i$, from which it follows that $y_i = y_j$ if and only if $i = j$. By $(\mathcal{K}1)$, it follows that $K_{w_1}^{y_i, 0, t} \cap K_{w_1}^{y_j, 0, t} \neq \emptyset$ if and only if $i = j$. By definition of \mathcal{A}_t , we have that

$$K \supseteq \biguplus_{i=1}^{\infty} K_{w_1}^{y_i, 0, t}.$$

If \mathcal{D}_t was empty then by Lemma 2.16 we would have $K = \biguplus_{i=1}^{\infty} K_{w_1}^{y_i, 0, t}$, expressing K as a disjoint union of a countable infinity of its subcomplexes. Hence $\mathcal{D}_t \neq \emptyset$, which proves the (contrapositive of the) stated result. ■

4.2. *Dust and GWVEs.* For all $t > 0$ we define

$$\mathcal{B}_n^t = \{w \in W_n; \nexists (s', w', p') \in \Pi_{(0, t]} \text{ such that } K_w \subseteq K_{w'}\}$$

and set $B_n^t = |\mathcal{B}_n^t|$. Note that this agrees with our preliminary definition from Section 1.5.

LEMMA 4.5. *For each fixed $t > 0$, $(B_n^t)_{n \geq 0}$ is a Galton-Watson process in a varying environment. The offspring distribution of individuals in stage $n \in \mathbb{N} \cup \{0\}$ is binomial with $|S|$ trials and success probability $e^{-tr_{n+1}}$. The initial distribution is a Bernoulli random variable with success probability e^{-tr_0} .*

PROOF. Consider B_n^t for some $n \in \mathbb{N} \cup \{0\}$, and in particular consider some $w \in W_n$ such that for all $(s, w, p) \in \Pi_{(0, t]}$, $K_w \not\subseteq K_{w'}$. We consider w to be an ‘individual at stage n ’ in the usual informal description of Galton-Watson processes.

By definition, $w \in \mathcal{B}_n^t$ if and only if $\{(s', w', p') \in \Pi_{(0, t]}; K_w \subseteq K_{w'}\}$ is empty. By definition of Π , for each $i \in S$ the probability that $\{(s', w', p') \in \Pi_{(0, t]}; K_{w_i} \subseteq K_{w'}\}$ is empty, conditional on $\{(s', w', p') \in \Pi_{(0, t]}; K_w \subseteq K_{w'}\} = \emptyset$, is the probability that $\Pi_{(0, t] \times \{w_i\}}$ is empty, which is $e^{-r_{n+1}t}$. By definition of Π , the sets $\Pi_{(0, t] \times \{w_i\}}$ and $\Pi_{(0, t] \times \{w_j\}}$ are independent for $i \neq j$. Hence B_n^t is a GWVE as required.

If $\Pi_{(0,t] \times W_0}$ is empty then the empty word \emptyset is an element of \mathcal{B}_0^t , otherwise it is not, and this occurs with probability $e^{-r_0 t}$. Thus we have the initial state as claimed. ■

LEMMA 4.6. *For each $t > 0$, almost surely, $\mathcal{D}_t = \bigcap_{n \in \mathbb{N}} \bigcup_{w \in \mathcal{B}_n^t} K_w$.*

PROOF. Consider $\omega \in \mathcal{A}$, and recall $\mathbb{P}[\mathcal{A}] = 1$ by Lemma 2.11. Fix $t > 0$.

Suppose first that $x \in \mathcal{D}_t$. Then for all $w \in W_*$ such that $x \in K_w$, $\Pi_{(0,t] \times \{w\}} = \emptyset$. If $x \in K_w$ and $K_w \subseteq K_{w'}$ then $x \in K_{w'}$ and $\Pi_{(0,t] \times \{w'\}} = \emptyset$. Hence, if $x \in K_w$ then $w \in \mathcal{B}_{|w|}^t$.

By (K1), for all $n \in \mathbb{N}$ there is precisely one $w \in W_n$ such that $x \in K_w$. Thus $x \in \bigcup_{w \in \mathcal{B}_n^t} K_w$ for all n , and hence $\mathcal{D}_t \subseteq \bigcap_{n \in \mathbb{N}} \bigcup_{w \in \mathcal{B}_n^t} K_w$.

Now suppose that $x \in \bigcap_{n \in \mathbb{N}} \bigcup_{w \in \mathcal{B}_n^t} K_w$. The argument above essentially works in reverse. Then $x \in \bigcup_{w \in \mathcal{B}_n^t} K_w$ for all n , from which it follows that if $x \in K_w$ we must have $\Pi_{(0,t] \times \{w\}} = \emptyset$. Thus $(E_m^{x,0,t}) = \emptyset$ and $x \in \mathcal{D}_t$. ■

LEMMA 4.7. *For all $t > 0$, $\mathbb{P}[\mathcal{D}_t = \emptyset] = \mathbb{P}[\exists n \in \mathbb{N}, B_n^t = 0]$.*

PROOF. Fix $t > 0$. Suppose first that for some (random) $n \in \mathbb{N}$, $B_n^t = 0$. By Lemma 4.6, $\mathcal{D}_t = \emptyset$.

Conversely, suppose that $\mathcal{D}_t = \emptyset$. By Lemma 4.4, \mathcal{A}_t is finite. Write $\mathcal{A}_t = \{x_1, \dots, x_N\}$. The same argument as in the final paragraph of the proof of Lemma 4.4 shows that, writing $\{K_{w(1)}, \dots, K_{w(M)}\} = \{X_{0,t}^{-1}(x_i); i = 1, \dots, N\}$ (relabelling with repeats, where $w(i) \in W_*$), $K = \biguplus_{i=1}^M K_{w(i)}$.

Let $n = \max\{|w(i)|; i = 1, \dots, M\}$ and consider some $w \in W_n$. By (K1) there is some $K_{w(i)}$ such that $K_w \subseteq K_{w(i)}$. Since $K_{w(i)} = X_{0,t}^{-1}(x_j)$, for some $j = 1, \dots, N$, it must be that $w(i) \notin \mathcal{B}_{|w(i)|}^t$. Hence $w \notin \mathcal{B}_n^t$. Since $w \in W_n$ was arbitrary, $B_n^t = 0$. ■

LEMMA 4.8. *For all $t > 0$, $\mathbb{P}[\lim_n B_n^t = \infty \text{ or } \exists n \geq \mathcal{N}, B_n^t = 0] = 1$. Further,*

$$(4.1) \quad \mathbb{P}[\mathcal{D}_t \neq \emptyset] = \mathbb{P}\left[\lim_{n \rightarrow \infty} B_n^t = \infty\right].$$

PROOF. To prove the first statement, we use the result of Theorem 1 in Jagers [1974], which is a restatement (with minor correction) of a result in Church [1967].

The probability of a individual at stage n in the process B_t having exactly one offspring is given by $p_{n1}^t = |S|e^{-r_n t}(1 - e^{-r_n t})^{|S|-1}$. Note that for $a \in [0, 1]$ and $n \geq 1$, $a(1 - a)^n \leq \frac{1}{n+1}(1 - \frac{1}{n+1})^n$. Since $|S| \geq 2$ we have $p_{n1}^t \leq$

$(1 - 1/|S|)^{|S|-1} < 1$. Hence $\sum_n (1 - p_{n1}^t) = \infty$, and from Jagers [1974] we have $\mathbb{P} \left[\lim_{n \rightarrow \infty} B_n^t = \infty \text{ or } \exists n \geq \mathcal{N}, B_n^t = 0 \right] = 1$. It follows immediately from this and Lemma 4.7 that $\mathbb{P} \left[\lim_{n \rightarrow \infty} B_n^t = \infty \right] = \mathbb{P} [\mathcal{D}_t \neq \emptyset]$. ■

LEMMA 4.9. *If $\sum r_n < \infty$ then $\mathbb{P} [\mathcal{D}_t = \emptyset \text{ or } \lambda^*(\mathcal{D}_t) > 0] = 1$.*

PROOF. The process $n \mapsto B_n^t / \mathbb{E} [B_n^t]$ is a discrete parameter, non-negative martingale. By the martingale convergence theorem there is some random variable L^t such that $\frac{B_n^t}{\mathbb{E} [B_n^t]} \rightarrow L^t$ almost surely.

Recall that in (1.7) we gave a formula for $\mathbb{E} [B_n^t]$. Since $\sum r_n < \infty$,

$$(4.2) \quad \mathbb{E} [B_n^t] \geq |S|^n \left(\mathbb{E} [B_0^t] \exp \left(-t \sum_1^\infty r_j \right) \right) = C |S|^n.$$

where $C = C(t) > 0$. In the language of Biggins and D'Souza [1992], (4.2) means that B^t is uniformly supercritical. Since the offspring distribution of B^t is uniformly bounded (by $|S|$), Theorem 2 of Biggins and D'Souza [1992] applies. In our notation this means that

$$(4.3) \quad \{B_n^t \rightarrow \infty\} = \{L^t > 0\}.$$

Now, suppose $\omega \in \mathcal{A}$ and that $\mathcal{D}_t \neq \emptyset$. By Lemma 4.6, for all $n \in \mathbb{N}$ we have $B_n^t = |\mathcal{B}_n^t| \geq 1$. By the first part of Lemma 4.8 it follows that (almost surely) $B_n^t \rightarrow \infty$ as $n \rightarrow \infty$. By (4.3), $\lim_{n \rightarrow \infty} \frac{B_n^t}{\mathbb{E} [B_n^t]} > 0$. From this and (4.2),

$$(4.4) \quad \liminf_{n \rightarrow \infty} \frac{B_n^t}{C |S|^n} > 0$$

where the \liminf could potentially be infinite. In fact, though, $B_n^t \leq B_0^t |S|^n \leq |S|^n$ so the \liminf in (4.4) is finite. We write $l = \liminf_{n \rightarrow \infty} \frac{B_n^t}{C |S|^n}$ where $l \in (0, \infty)$ (note l is random). Then there exists $N \in \mathbb{N}$ such that for all $n > N$, $\frac{B_n^t}{C |S|^n} \geq l/2$. So for all $n > N$ we have $B_n^t \geq \frac{Cl}{2} |S|^n$.

Note that the sets $\bigcup_{w \in \mathcal{B}_n^t} K_w$ are decreasing as n increases. By Lemma 4.6, $\lambda(\mathcal{D}_t) = \lim_n \lambda \left(\bigcup_{w \in \mathcal{B}_n^t} K_w \right)$. Recall that $\lambda(K_w) > 0$ and $\lambda(K_w)$ depends only on $|w|$ (by Lemma 2.2 and (K3) respectively). Hence $\lambda(K_w) = \frac{\lambda(K)}{|S|^n}$. By (K1),

$$\lambda \left(\bigcup_{w \in \mathcal{B}_n^t} K_w \right) = \sum_{w \in \mathcal{B}_n^t} \lambda(K_w) = |\mathcal{B}_n^t| \lambda(K_w) = \frac{B_n^t}{|S|^n} \lambda(K).$$

Thus, for $n > N$, $\lambda\left(\bigcup_{w \in \mathcal{B}_n^t} K_w\right) \geq Cl\lambda(K)$. Hence, $\lambda(\mathcal{D}_t) > 0$. ■

4.3. *Some real analysis.* To apply Lemmas 4.7 and 4.8 we need a characterisation of the extinction/explosion of B^t . A general degeneracy criterion for Galton-Watson processes in varying environments seems not to be known. However, our offspring distributions are binomial and uniformly bounded above, and in this case (amongst others) necessary and sufficient conditions can be found in Jirina [1976]. Recall that $\mathbb{E}[B_n^t] = \mathbb{E}[B_0^t] m_n^t$ where

$$m_n^t = \exp\left(n \log |S| - t \sum_{j=1}^n r_j\right).$$

Clearly $m_n^t \in (0, \infty)$. The quantity $\inf_n m_n^t$ will feature in these conditions, along with

$$g^t = \sum_{n=1}^{\infty} \frac{(1 - e^{-r_{n+1}t})|S| + |S|e^{-r_{n+1}t} - 1}{|S|e^{-r_{n+1}t} m_n^t}.$$

For all $x > 0$ and $n \in \mathbb{N}$, $n \geq 2$, it holds that $(1 - x)^n + nx - 1 > 0$, and hence $g^t \in (0, \infty]$.

The following result comes from applying Lemma 1.1 and Theorem 2.3 of Jirina [1976].

LEMMA 4.10. $\mathbb{P}[\exists n \in \mathbb{N}, B_n^t = 0] < 1$ if and only if both $\inf m_n^t > 0$ and $g^t < \infty$.

The remainder of Section 4.3 will consist of entirely real analysis, with the objective of pinning down the behaviour of $\inf_n m_n^t$ and g^t .

LEMMA 4.11. Suppose that $v \in (0, \infty)$ is such that $\inf_n m_n^v > 0$. Then for all $u \in (0, v)$, $\inf_n m_n^u > 0$ and $g^u < \infty$.

PROOF. Let $\inf_n m_n^v > 0$. Suppose that $\epsilon > 0$ and for infinitely many n we have $\frac{1}{n} \sum_{j=1}^n r_j > \frac{\epsilon + \log |S|}{v}$. For such n ,

$$\begin{aligned} m_n^v &= \left(\exp\left(\log |S| - v \frac{1}{n} \sum_{j=1}^n r_j\right) \right)^n \\ &\leq \left(\exp\left(\log |S| - v \frac{\epsilon + \log |S|}{v}\right) \right)^n = e^{-\epsilon n}. \end{aligned}$$

This may not occur for infinitely many n since $\inf m_n^v > 0$.

So, we may assume that both $\inf_n m_n^v > 0$, and

$$(4.5) \quad \limsup_n \frac{1}{n} \sum_1^n r_j \leq \frac{\log |S|}{v}.$$

Let $u \in (0, v)$. Then, for some $\epsilon > 0$ we have $0 < u \limsup_n \frac{1}{n} \sum_1^n r_j < \log |S| - \epsilon$. Hence, there exists $N \in \mathbb{N}$ (dependent on ϵ) such that for all $n > N$, $0 < u \frac{1}{n} \sum_1^n r_j < \log |S| - \epsilon$. Thus,

$$m_n^u = \left(\exp \left(\log |S| - u \frac{1}{n} \sum_1^n r_j \right) \right)^n \geq (\exp(\epsilon))^n,$$

so clearly $\inf_n m_n^u > 0$. Also

$$g^u \leq \sum_1^\infty \frac{|S|}{|S| e^{-r_{n+1} u} m_n^u} = |S| \sum_1^\infty \frac{1}{m_{n+1}^u} \leq |S| \sum_1^\infty \frac{1}{(e^\epsilon)^{n+1}} < \infty$$

as required. ■

LEMMA 4.12. *The following hold:*

1. $\inf_n m_n^t = 0$ for all $t > 0$ if and only if $\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_1^n r_j = \infty$.
2. $\inf_n m_n^t > 0$ for all $t > 0$ if and only if $\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_1^n r_j = 0$.
3. $\exists t_0 \in (0, \infty)$ such that $t < t_0 \Rightarrow \inf_n m_n^t > 0$ and $t > t_0 \Rightarrow \inf_n m_n^t = 0$, if and only if, $\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_1^n r_j \in (0, \infty)$.

PROOF. We give each case in turn.

CASE 1: Suppose that $\limsup_n \frac{1}{n} \sum_1^n r_j = \infty$. For any $t > 0$, we can pick a subsequence (r_{i_n}) of (r_n) such that for all n , $\frac{1}{i_n} \sum_1^{i_n} r_j > \frac{\log |S| + 1}{t}$. Hence $m_{i_n}^t \leq (\exp(-1))^{i_n}$ for all n , and since $i_n \rightarrow \infty$ it follows that $\inf_n m_n^t = 0$.

Conversely, if $\limsup_n \frac{1}{n} \sum_1^n r_j = C < \infty$ then for $t = \frac{\log |S|}{2C} > 0$ we note that

$$m_n^t = \left(\exp \left(\log |S| - \frac{\log |S|}{2} \frac{1}{C} \frac{1}{n} \sum_1^n r_j \right) \right)^n.$$

For sufficiently large n , $\frac{1}{n} \sum_1^n r_j \leq \frac{3}{2}C$, and hence for sufficiently large n , $m_n^t \geq (\exp(\frac{1}{3} \log |S|))^n$. Hence $\inf_n m_n^t > 0$.

CASE 2: Suppose that $\limsup_n \frac{1}{n} \sum_1^n r_j = 0$ and let $t > 0$. Then, for all sufficiently large n we have $\frac{1}{n} \sum_1^n r_j \leq \frac{1}{2t}$. Hence, for all sufficiently large n , $m_n^t \geq (\exp(\frac{1}{2} \log |S|))^n$. Thus $\inf_n m_n^t > 0$.

Conversely, suppose that $\inf_n m_n^t > 0$ for all t . Fixing t , and using the first step of the proof of Lemma 4.11, we obtain from (4.5) that $\limsup_n \frac{1}{n} \sum_1^n r_j \leq \frac{\log |S|}{t}$. However, we have $\inf_n m_n^t > 0$ for all $t > 0$, so $\limsup_n \frac{1}{n} \sum_1^n r_j = 0$.

CASE 3: It follows immediately from the definition of m_n^t that for $t < s$, $m_n^t \leq m_n^s$. Thus, $\inf_n m_n^t \leq \inf_n m_n^s$. Hence, for each $|S|$ and (r_n) there is a unique $t^* \in [0, \infty]$ such that $\inf_n m_n^t > 0$ for $t < t^*$ and $\inf_n m_n^t = 0$ for $t > t^*$. This case now follows from cases 1 and 2. ■

LEMMA 4.13. *There exist $C_1, C_2 \in (0, \infty)$ (dependent only upon $|S|$) such that for all $t > 0$,*

$$C_1 \sum_n \frac{e^{-r_{n+1}t}}{m_n^t} \leq g^t \leq C_2 \sum_n \frac{e^{-r_{n+1}t}}{m_n^t}.$$

PROOF. Let $f_n : (0, \infty) \rightarrow (0, \infty)$ be defined by $f_n(x) = (1-x)^n + nx - 1$. It is elementary to show that there exists $C_1, C_2 \in (0, \infty)$ (dependent only upon $|S|$) such that for all $x \in (0, 1]$, $C_1 x^2 \leq f_{|S|}(x) \leq C_2 x^2$. Since

$$g^t = \sum_n \frac{f_{|S|}(e^{-r_{n+1}t})}{|S| e^{-r_{n+1}t} m_n^t},$$

the stated result follows. ■

4.4. Proof of Theorem 2 and Corollaries 3-5.

PROOF. [of Theorem 2 and Corollary 3] It is elementary to show that our criteria assign each possible choice of $|S|$ and (r_n) to precisely one phase. Hence, it suffices to show that the given criteria for each phase are sufficient.

Recall that for at least one n , $r_n > 0$. Since $K = \bigcup_{|w|=n} K_w$, $\mathbb{P}[\mathcal{P}_1^t] > 0$ for all $t > 0$, regardless of the phase. Recall the meaning of $\mathbb{P}\langle A * B \rangle$ from Definition 2.17.

LOWER SUBCRITICAL: Suppose that $\sum |S|^n < \infty$ and let $t > 0$. Then the total rate of Π is finite, so $\Pi_{(0,t]}$ is almost surely a finite set. By Lemma 2.16, $\mathcal{A}(\lambda_t^*)$ is almost surely finite.

Let $W(1) = \{w \in W_*; K_w \subseteq K_1\}$. Since the total rate of Π is finite, the rates of events occurring inside K_1 is finite. Since $r_1 > 0$,

$$\mathbb{P}[\Pi_{(0,t] \times \{\emptyset\}} = \emptyset, \Pi_{(0,t] \times W(1)} = \emptyset, \Pi_{(0,t] \times \{2\}} \neq \emptyset] > 0.$$

Thus, with positive probability we have $K_1 \subseteq \mathcal{D}_t$ and a non trivial block K_2 . Thus $\mathbb{P}[\mathcal{P}_2^t] > 0$.

Since $\sum |S|^n r_n < \infty$, we have also that $\sum r_n < \infty$. By Lemma 4.9, if $\mathcal{D}_t \neq \emptyset$ then (a.s.) $\lambda_t^*(\mathcal{D}_t) > 0$. Since $\mathcal{A}(\lambda_t^*)$ is almost surely finite, $\mathbb{P}\langle \mathcal{P}_1^t * \mathcal{P}_2^t \rangle = 1$.

UPPER SUBCRITICAL: Suppose $\sum |S|^n = \infty$ and $\sum r_n < \infty$. By Lemma 4.1, $\lambda_t^*(\mathcal{D}_t) > 0$ with positive probability. If $\lambda_t^*(\mathcal{D}_t) > 0$ and $\mathcal{A}(\lambda_t^*)$ was finite (with positive probability), then $\mathcal{D}_t \neq \emptyset$ and by Lemma 4.2 it would follow that \mathcal{D}_t contained a subcomplex, contradicting Lemma 4.3. Hence, with positive probability we have that $\lambda_t^*(\mathcal{D}_t) > 0$ and $\mathcal{A}(\lambda_t^*)$ is infinite. Since λ_t^* is a finite measure it can have at most countably many atoms. Hence $\mathbb{P}[\mathcal{P}_3^t] > 0$.

If $\mathcal{D}_t \neq \emptyset$ then Lemma 4.9 shows that (a.s.) $\lambda_t^*(\mathcal{D}_t) > 0$. If $\mathcal{A}(\lambda_t^*)$ is finite then by Lemma 4.3 either $\mathcal{D}_t = \emptyset$ or \mathcal{D}_t contains a subcomplex of K . By Lemma 4.2, this second case does not occur (almost surely). Hence in this case $\mathcal{A}(\lambda_t^*)$ must be infinite. Since λ_t^* is a finite measure, $\mathcal{A}(\lambda_t^*)$ is countably infinite. Hence $\mathbb{P}\langle \mathcal{P}_1^t * \mathcal{P}_3^t \rangle = 1$.

SEMICRITICAL: Suppose that $\sum r_n = \infty$, and that $\limsup_n \frac{1}{n} \sum_1^n r_j = 0$. Then, by Lemma 4.12, for all $t > 0$ we have $\inf_n m_n^t > 0$. Hence, by Lemma 4.11, for all $t > 0$ we have $g^t < \infty$.

By Lemma 4.10, $\mathbb{P}[\exists n \geq \mathcal{N}, B_n^t = 0] < 1$, so by Lemma 4.8 we have $\mathbb{P}[B_n^t \rightarrow \infty] > 0$ and $\mathbb{P}[\mathcal{D}_t \neq \emptyset] > 0$. If $\mathcal{D}_t \neq \emptyset$ and $\mathcal{A}(\lambda_t^*)$ is finite then by Lemma 4.2 we must have $\lambda_t^*(\mathcal{D}_t) > 0$, but this does not occur almost surely by Lemma 4.1. Hence, $\mathbb{P}[\mathcal{D}_t \neq \emptyset, \mathcal{A}(\lambda_t^*) \text{ is countably infinite}] > 0$. By Lemma 4.1, $\mathbb{P}[\lambda_t^*(\mathcal{D}_t) = 0] = 1$, and hence we have $\mathbb{P}[\mathcal{P}_4^t] > 0$.

Lemma 4.1 implies that if \mathcal{D}_t is non-empty then almost surely we have $\lambda_t^*(\mathcal{D}_t) = 0$ and $\mathcal{D}_t \neq \emptyset$. In this case, if $\mathcal{A}(\lambda_t^*)$ was finite, Lemma 4.2 would imply that $\lambda_t^*(\mathcal{D}_t) > 0$, which is a contradiction, so in fact $\mathcal{A}(\lambda_t^*)$ must be infinite. Since λ_t^* is a finite measure, $\mathcal{A}(\lambda_t^*)$ must be countably infinite. Hence $\mathbb{P}\langle \mathcal{P}_1^t * \mathcal{P}_4^t \rangle = 1$.

SUPERCRITICAL: Suppose that $\limsup_n \frac{1}{n} \sum_1^n r_j = \infty$. It follows immediately that $\sum r_n = \infty$. By Lemma 4.12, for all $t > 0$ we have $\inf_n m_n^t = 0$. By Lemma 4.10, $\mathbb{P}[\exists n \geq \mathcal{N}, B_n^t = 0] = 1$, and it follows immediately by Lemma 4.7 that $\mathbb{P}[\mathcal{D}_t = \emptyset] = 1$. By Lemma 4.4, $\mathbb{P}[\mathcal{P}_1^t] = 1$.

CRITICAL: Suppose that $\limsup_n \frac{1}{n} \sum_1^n r_j \in (0, \infty)$. It follows immediately that $\sum r_n = \infty$. Recall that earlier in this proof we showed that $t \mapsto \inf m_n^t$ decreases as t increases. By Lemma 4.12, there is some $t_0 \in (0, \infty)$ such that $t \in (0, t_0)$ (where $t_0 \in (0, \infty)$) we have $\inf m_n^t > 0$ and that for $t \in (t_0, \infty)$ we have $\inf m_n^t = 0$.

By Lemma 4.11 for $t \in (0, t_0)$ we have also that $g^t < \infty$, and as in the semicritical case we have that $\mathbb{P}\langle \mathcal{P}_1^t * \mathcal{P}_4^t \rangle = 1$. As in the supercritical case, for $t \in (t_0, \infty)$ we have $\mathbb{P}[\mathcal{P}_1^t] = 1$.

This completes the proof of Theorem 2. Corollary 3 follows immediately

from Theorem 2 and the definition of the phases. ■

PROOF. [of Corollary 4] Let $\mathcal{L} = \limsup_n \frac{1}{n} \sum_1^n r_j$ and let $t^* = \frac{\log |S|}{\mathcal{L}}$. and suppose that $0 < \mathcal{L} < \infty$. By Theorem 2, the critical time t_0 is given by $t_0 = \sup\{t > 0; \mathbb{P}[\mathcal{D}_t \neq \emptyset] > 0\}$.

Consider first when $t < t^*$. Then there exists $\epsilon \in (0, \log |S|)$ such that $t \leq \frac{\log |S| - \epsilon}{\mathcal{L}}$. Hence,

$$m_n^t \geq \left(\exp \left(\log |S| - \frac{\log |S| - \epsilon}{\mathcal{L}} \frac{1}{n} \sum_1^n r_j \right) \right)^n.$$

There exists $N \in \mathbb{N}$ such that for all $n > N$, $\frac{1}{n} \sum_1^n r_j \leq \frac{\log |S| - \epsilon/2}{\log |S| - \epsilon} \mathcal{L}$. Hence, for all $n > N$,

$$m_n^t \geq \left(\exp \left(\log |S| - (\log |S| - \epsilon) \frac{\log |S| - \epsilon/2}{(\log |S| - \epsilon)} \right) \right)^n = (\exp(\epsilon/2))^n$$

Thus $m_n^t \rightarrow \infty$ and hence $\inf m_n^t > 0$. By Lemma 4.11, we have $\inf_n m_n^t > 0$ and $g^t < \infty$ for all $t \in (0, t^*)$. By Lemmas 4.8 and 4.10, $\mathbb{P}[\mathcal{D}_t \neq \emptyset] > 0$ for $t < t^*$.

Now consider $t > t^*$. For some $\epsilon > 0$, $\frac{\log |S| + \epsilon}{\mathcal{L}} \leq t$. There exists a subsequence (r_{i_n}) of (r_n) such that for all n , $\lim_{n \rightarrow \infty} \frac{1}{i_n} \sum_1^{i_n} r_j = \mathcal{L}$. Hence

$$m_{i_n}^t \leq \left(\exp \left(\log |S| - \frac{\log |S| + \epsilon}{\mathcal{L}} \frac{1}{i_n} \sum_1^{i_n} r_j \right) \right)^{i_n}.$$

Therefore, there exists $N \in \mathbb{N}$ such that for all $n > N$,

$$\left| \frac{1}{i_n} \sum_1^{i_n} r_j - \mathcal{L} \right| \leq \mathcal{L} \frac{\epsilon/2}{\log |S| + \epsilon}.$$

For all such n , $\left| \frac{1}{\mathcal{L}} \frac{1}{i_n} \sum_1^{i_n} r_j - 1 \right| \leq \frac{\epsilon/2}{\log |S| + \epsilon}$, so as

$$\frac{1}{\mathcal{L}} \frac{1}{i_n} \sum_1^{i_n} r_j \geq 1 - \frac{\epsilon/2}{\log |S| + \epsilon} = \frac{\log |S| + \epsilon/2}{\log |S| + \epsilon}.$$

Hence, for all $n > N$,

$$m_{i_n}^t \leq \left(\exp \left(\log |S| - (\log |S| + \epsilon) \frac{\log |S| + \epsilon/2}{\log |S| + \epsilon} \right) \right)^{i_n} = (\exp(-\epsilon/2))^{i_n}.$$

Thus $m_{i_n}^t \rightarrow 0$ and $\inf m_n^t = 0$. By Lemmas 4.8 and 4.10, $\mathbb{P}[\mathcal{D}_t = \emptyset] = 0$ for $t > t^*$.

Combining the two cases, we have that $t_0 = t^*$. ■

PROOF. [of Corollary 5] Since X is critical, $\mathcal{L} = \limsup_n \frac{1}{n} \sum_1^n r_j \in (0, \infty)$. By Corollary 4, $t_0 = \frac{\log |S|}{\mathcal{L}}$. Let $\mathcal{L} = \limsup_n \frac{1}{n} \sum_1^n r_j$, and suppose $\mathcal{L} \in (0, \infty)$. Define $a_n \in \mathbb{R}$ by $r_n = \mathcal{L} + a_n$. In this notation, $m_n^{t_0} = \exp(-t_0 \sum_1^n a_j)$

We consider two cases. If $\limsup_n \sum_1^n a_j = \infty$ then $\inf_n m_n^{t_0} = 0$. The same argument as used for the supercritical case in Theorem 2 then shows that $\mathbb{P}[\mathcal{P}_1^{t_0}] = 1$.

If $\limsup \sum_1^n a_j < \infty$ then $\inf m_n^{t_0} > 0$. By Lemma 4.13 there exists $C \in (0, \infty)$ such that $g^{t_0} \geq \frac{C}{\inf_n m_n^{t_0}} \sum_1^n e^{-r_{n+1}t_0}$. Since $\limsup \frac{1}{n} \sum_1^n r_j \in (0, \infty)$, (r_n) has a subsequence (r_{i_n}) such that $\limsup_n r_{i_n} < \infty$. Hence $g^{t_0} = \infty$. By Lemma 4.10, $\mathbb{P}[\exists n \in \mathbb{N}, B_n^{t_0} = 0] = 1$, and hence $\mathbb{P}[\mathcal{D}_{t_0} = \emptyset] = 1$. By Lemma 4.4, $\mathbb{P}[\mathcal{P}_1^{t_0}] = 1$. ■

5. The Hausdorff dimension of the dust. Let $\|\cdot\|$ denote the Euclidean norm on \mathbb{R}^d , and let \mathcal{L}^d denote d dimensional Lebesgue measure. Let A° denote the (topological) interior of the set A , and let the diameter of A be given by $\text{diam}(A) = \sup\{\|x - y\|; x, y \in A\}$. Recall that a similarity f is a function between subsets of \mathbb{R}^d such that for some $\eta \in (0, \infty)$ and all x, y , $\|f(x) - f(y)\| = \eta\|x - y\|$. We write $\eta = \text{lip}(f)$. Recall also that $\dim_{\mathcal{H}}(A)$ denotes the Hausdorff dimension of A (for $A \subseteq \mathbb{R}^d$ this is with respect to the metric $(x, y) \mapsto \|x - y\|$).

In order to link our results to those of Durand [2009], we must impose some extra assumptions on K .

DEFINITION 5.1. We say K is D -compatible if $K \subseteq \mathbb{R}^d$, $D_K = \|\cdot\|$, and

1. For all $w \in W_*$, K_w is compact.
2. For all $w \in W_*$ and $i \in S$ there exists a similarity $f^{(w,i)} : K_w \rightarrow K_{wi}$.
3. There exists $\epsilon, \epsilon' \in (0, 1)$ and a sequence $(l_n) \subseteq [\epsilon, \epsilon']$ such that for all $w \in W$, $\text{lip}(f^{(w,i)}) = l_{|w|}$.
4. There exists $\kappa > 0$ such that for all $w \in W_*$, $\mathcal{L}^d(K_w^\circ) \geq \kappa \text{diam}(K_w)^d$.

In words, $l_{|w|}$ is the scale factor which K_w must be shrunk by to get a copy of K_{wi} . Condition 3 says that this depends only on w , and not on i .

THEOREM 5.2. Suppose that K is D -compatible and that $\mathbb{P}[\mathcal{D}_t \neq \emptyset] > 0$. Let $\mathcal{L} = \limsup_n \frac{1}{n} \sum_1^n r_j$ and $\mathcal{S} = \limsup_n \frac{1}{n} \sum_1^n (-\log l_n)$. Conditional on $\{\mathcal{D}_t \neq \emptyset\}$,

$$\dim_{\mathcal{H}}(\mathcal{D}_t) = \left(\frac{\log |S| - t\mathcal{L}}{\mathcal{S}} \right) \vee 0.$$

By setting $r_n = 0$, it follows that $\dim_{\mathcal{H}}(K) = \frac{\log |S|}{\mathcal{S}}$.

PROOF. For each $s > 0$ and $n \in \mathbb{N}$ let $\alpha_{s,n}^t = |S|(l_n)^s e^{-r_{n+1}t}$ and $\rho^t(s) = \liminf_n \frac{1}{n} \sum_{j=1}^n \log \alpha_{s,j}$. Using (K1) and the D-compatibility conditions we apply Theorem 1 of Durand [2009], which yields that, if $\mathcal{D}_t \neq \emptyset$, then $\dim_{\mathcal{H}}(\mathcal{D}_t) = \sup\{s \in [0, \infty); \rho^t(s) > 0\}$.

By Theorem 2, if $\mathbb{P}[\mathcal{D}_t \neq \emptyset] > 0$ then $0 \leq \limsup_n \frac{1}{n} \sum_1^n r_j < \infty$. Note also that by 3 of the D-compatibility conditions, $0 \leq -\log \epsilon' \leq -\log(l_n) \leq -\log \epsilon < \infty$. A short calculation shows that

$$\rho^t(s) = \log |S| - t \limsup_{n \rightarrow \infty} \left(\frac{1}{n} \sum_1^n r_j \right) - s \limsup_{n \rightarrow \infty} \left(\frac{1}{n} \sum_1^n (-\log l_n) \right).$$

The result follows. ■

References.

- O. Angel, N. Berestycki, and V. Limic. Global divergence of spatial coalescents. *Probab. Th. Rel. Fields*, Published online:1–45, 2010.
- R. Arratia. Coalescing Brownian motions on the line. *Ph. D. Thesis, University of Wisconsin, Madison*, 1979.
- N. H. Barton, A. M. Etheridge, and A. Véber. A new model for evolution in a spatial continuum. *Electron. J. Probab.*, 15:162–216, 2010a.
- N. H. Barton, J. Kelleher, and A. M. Etheridge. A new model for large scale population dynamics: quantifying phylogeography. *Evolution*, 64:2701–2715, 2010b.
- J. Berestycki, N. Berestycki, and J. Schweinsberg. Small-time behavior of β -coalescents. *Ann. Inst. H. Poincaré Probab. Statist.*, 44(2):214–238, 2008.
- J. Berestycki, N. Berestycki, and V. Limic. The Λ -coalescent speed of coming down from infinity. *Ann. Probab.*, 38:207–233, 2010.
- N. Berestycki. *Recent Progress In Coalescent Theory*, volume 16. Ensaio Matemáticos, 2009.
- N. Berestycki, A. M. Etheridge, and M. Hutzenthaler. Survival, extinction and ergodicity in a spatially continuous population model. *Markov Process. Rel. Fields*, 15(3):265–288, 2009.
- J. Bertoin and J. F. Le Gall. The Bolthausen-Sznitman coalescent and the genealogy of continuous-state branching processes. *Probab. Th. Rel. Fields*, 117:249–266, 2000.
- J. Bertoin and J. F. Le Gall. Stochastic flows associated to coalescent processes. *Probab. Th. Rel. Fields*, 126:261–288, 2003.
- J. Bertoin and J. F. Le Gall. Stochastic flows associated to coalescent processes III. Limit theorems. *Illinois J. Math.*, 50:147–181, 2006.
- J. Biggins and J. D’Souza. The supercritical Galton-Watson process in varying environments. *Stochastic Process. Appl.*, 42:39–47, 1992.
- M. Birkner, J. Blath, M. Capaldo, A. M. Etheridge, M. Möhle, J. Schweinsberg, and A. Wakolbinger. α -stable branching and β -coalescents. *Electron. J. Probab.*, 10(9):303–325, 2005.
- E. Bolthausen and A. Sznitman. On Ruelle’s probability cascades and an abstract cavity method. *Comm. Math. Phys.*, 197:247–276, 1998.
- J. D. Church. Composition limit theorems for probability generating functions. *MRC Techn. Summ. Rept.*, (732):1–41, 1967.

- D. A. Dawson and A. Greven. Hierarchically interacting Fleming-Viot processes with selection and mutation: Multiple space time scale analysis and quasi-equilibria. *Electron. J. Probab.*, 4:1–81, 1999.
- P. Donnelly and T. G. Kurtz. Genealogical processes for Fleming-Viot models with selection and recombination. *Ann. Appl. Probab.*, 9(4):1091–1148, 1999.
- A. Durand. Random fractals and tree-indexed Markov chains. *Rev. Mat. Iberoam.*, 25(3):1089–1126, 2009.
- A. M. Etheridge. Drift, draft and structure: some mathematical models of evolution. *Banach Center Publ.*, 80:121–144, 2008.
- A. M. Etheridge and A. Véber. The spatial Λ -Fleming-Viot process on a large torus: genealogies in the presence of recombination. *arXiv:1106.4050v1*, 2011.
- K. Falconer. *Fractal Geometry - Mathematical Foundations and Applications (second edition)*. John Wiley, 2003.
- D. H. Fearn. Galton-Watson processes with generation dependence. *Proc. Sixth Berkeley Symp. on Math. Statist. and Prob.*, 4:159–172, 1972.
- N. Freeman. The number of non-singleton blocks in the Λ -coalescent. <http://arxiv.org/abs/1111.1660>, pages 1–21, 2011.
- A. Greven, V. Limic, and A. Winter. Representation theorems for interacting Moran models, interacting Fisher-Wright diffusions and applications. *Electron. J. Probab.*, 10(39):1286, 2005.
- J. E. Hutchinson. Fractals and self-similarity. *Indiana University Mathematics Journal*, 30(5):712–747, 1981.
- P. Jagers. Galton-Watson processes in varying environments. *J. Appl. Probab.*, 11:174–178, 1974.
- M. Jirina. Extinction of non-homogeneous Galton-Watson processes. *J. Appl. Probab.*, 13:132–137, 1976.
- J. F. C. Kingman. The coalescent. *Stochastic Process. Appl.*, 13:235–248, 1982.
- Y. Le Jan and O. Raimond. Flows, coalescence and noise. *Ann. Probab.*, 32(2):1247–1315, 2004.
- V. Limic. On the speed of coming down from infinity of Ξ -coalescent processes. *Electron. J. Probab.*, 15:217–240, 2010.
- V. Limic and A. Sturm. The spatial Λ -coalescent. *Electron. J. Probab.*, 11:363–393, 2006.
- M. Möhle and S. Sagitov. A classification of coalescent processes for haploid exchangeable population models. *Ann. Probab.*, 29(4):1547–1562, 2001.
- P. Mörters. *New perspectives in stochastic geometry*, chapter 8. Oxford University Press, 2010.
- J. Neveu. A continuous-state branching process in relation with the GREM model of spin glass theory. *Rapport interne no 267, Ecole Polytechnique*, 1992.
- J. Pitman. Coalescents with multiple collisions. *Ann. Probab.*, 27(4):1870–1902, 1999.
- S. Sagitov. The general coalescent with asynchronous mergers of ancestral lines. *J. Appl. Probab.*, 36:1116–1125, 1999.
- J. Schweinsberg. A necessary and sufficient condition for the Λ -coalescent to come down from infinity. *Electron. Comm. Probab.*, 5:1–11, 2000a.
- J. Schweinsberg. Coalescents with simultaneous multiple collisions. *Electron. J. Probab.*, 5:1–50, 2000b.